Progress in Probability, Vol. 66, 19–53

Slepian's Inequality, Modularity and Integral Orderings

J. Hoffmann-Jørgensen

Abstract. Slepian's inequality comes in many variants under different sets of regularity conditions. Unfortunately, some of these variants are wrong and other variants are imposing to strong regularity conditions. The first part of this paper contains a unified version of Slepian's inequality under minimal regularity conditions, covering all the variants I know about. It is well known that Slepian's inequality is closely connected to integral orderings in general and the supermodular ordering in particular. In the last part of the paper I explore this connection and corrects some results in the theory of integral orderings.

Mathematics Subject Classification (2010). Primary 60E15; Secondary 60A10.

Keywords. Integral orderings, modular functions, Gaussian vectors.

1. Introduction

Throughout this paper, we let $(\Omega, \mathcal{F}, P)$ denote a fixed probability space. If $k \geq 1$ is an integer, we set $[k] := \{1, \dots, k\}$. If $X = (X_1, \dots, X_k)$ is a random vector such that $X_1, \dots, X_k \in L^2(P)$, we let $\bar{X}_i := X_i - EX_i$ denote *the centered random variables* for $i \in [k]$ and we let $\Sigma^X = \{\sigma_{ij}^X\}$ and $\Pi^X = \{\pi_{ij}^X\}$ denote the *covariance matrix* and *intrinsic metric* of X; that is:

$$\sigma_{ij}^X := E(\bar{X}_i \bar{X}_j) \textit{ and } \pi_{ij}^X := E(\bar{X}_i - \bar{X}_j)^2 \quad \forall\, i, j \in [k].$$

Note that $\pi_{ij}^X = \sigma_{ii}^X + \sigma_{jj}^X - 2\,\sigma_{ij}^X$ for all $i, j \in [k]$ and that $d(i,j) := \sqrt{\pi_{ij}^X}$ is a Hilbertian pseudo-metric on $[k]$.

It is well known that Slepian's inequality is an important tool in the theory of Gaussian processes. Let $X = (X_1, \dots, X_k)$ and $Y = (Y_1, \dots, Y_k)$ be k-dimensional Gaussian vectors with zero means. Slepian's inequality comes in many variants; see [2, 6, 7, 9, 10, 12, 16, 18], but in essence it states that $Ef(Y) \leq Ef(X)$ for all

$f : \mathbf{R}^k \to \mathbf{R}$ satisfying

$$(\sigma_{ij}^X - \sigma_{ij}^Y)\tfrac{\partial^2 f}{\partial x_i \partial x_j}(x) \geq 0 \quad \text{or} \quad (\pi_{ij}^Y - \pi_{ij}^X)\tfrac{\partial^2 f}{\partial x_i \partial x_j}(x) \geq 0 \quad \forall i, j \tag{1.1}$$

plus some regularity conditions. Condition (1.1) indicates that f should be sufficiently smooth (at least twice differentiable), but Slepian's inequality is often used for indicator functions which are not even continuous. In most of the literature the indicator case and the smooth case are treated separately. The most general form of Slepian's inequality is found in [7] and [9] where (1.1) is interpreted in the sense of Schwartz distributions. However, Theorem 3.11 on p. 74 in [9] is false as it stands:

Example A: Let $k \geq 2$ be an integer, let $U, U_1, \ldots, U_k$ be independent $N(0,1)$-distributed random variables and set $X = (U, \ldots, U)$ and $Y = (U_1, \ldots, U_k)$. Then X and Y are Gaussian random vectors such that $\sigma_{ii}^Y = \sigma_{ii}^X$ and $\sigma_{ij}^Y = 0 < 1 = \sigma_{ij}^X$ for $1 \leq i \neq j \leq 1$. Let $D := \{x \in \mathbf{R}^k \mid x_1 = \cdots = x_k\}$ denote the diagonal in $\mathbf{R}^k$ and set $f = -1_D$. Since $f = 0$ Lebesgue a.e. we have $Ef(Y) = 0$ and $\frac{\partial^2 f}{\partial x_i \partial x_j} = 0$ in distribution sense for all $1 \leq i, j \leq k$ and since $P(X \in D) = 1$, we have $Ef(X) = -1$ showing that Theorem 3.11 on p. 74 in [9] fails in this case. Many other counterexamples can be constructed in a similar manner.

This observation calls for a closer glance at the validity of Slepian's inequality and Section 2 of this paper will be devoted to establish a unified form of Slepian's inequality under minimal regularity conditions on f.

Slepian's inequality is intimately connected with integral orderings in general and the supermodular ordering in particular. If S and T are sets, we let 2^S denote the set of all subsets of S, we let T^S denote the set of all functions from S into T, and we let $B(S)$ denote the set of all bounded, real-valued functions on S. Recall that $(T, \leq)$ is a *proset* if T is a non-empty set and $\leq$ is a relation on T such that $\leq$ is *reflexive* ($t \leq t \; \forall t \in T$) and *transitive* ($t \leq u \;,\; u \leq v \;\Rightarrow\; t \leq v$).

Let $(S, \mathcal{A})$ be a measurable space; that is, S is a non-empty set and $\mathcal{A}$ is a σ-algebra on S. Then we let $M(S, \mathcal{A})$ denote the set of all $\mathcal{A}$-measurable functions from S into $\mathbf{R}$ and we let $\Pr(S, \mathcal{A})$ denote the set of all probability measures on $(S, \mathcal{A})$. Let $\Phi \subseteq M(S, \mathcal{A})$ be a given set of functions. Then it is customary to define *the Φ-integral ordering* on $\Pr(S, \mathcal{A})$, denoted $\leq_\Phi$, as follows: $\mu \leq_\Phi \nu$ if and only if $\int_S \phi \, d\mu \leq \int_S \phi \, d\nu$ for all $\phi \in \Phi \cap L^1(\mu) \cap L^1(\nu)$. Then $\leq_\Phi$ is a relation on $\Pr(S, \mathcal{B})$ which is reflexive but not transitive and exhibits strange properties:

Example B: Let $S = \mathbf{R}$ and let $\mathcal{B}$ denote the Borel σ-algebra on $\mathbf{R}$. Let Φ denote the set of all increasing, convex functions $\phi : \mathbf{R} \to \mathbf{R}$. Let μ be a Borel probability measure such that $\int_{\mathbf{R}} x^+ \, \mu(dx) = \infty$. Then $\Phi \cap L^1(\mu)$ is the set of all constant functions and so we have $\mu \leq_\Phi \nu$ and $\nu \leq_\Phi \mu$ for all $\nu \in \Pr(\mathbf{R}, \mathcal{B})$. In particular, we see that $\leq_\Phi$ is not transitive and that the integral ordering $\leq_\Phi$ is *not* a preordering.

To avoid such peculiarities, I shall introduce a slight modification of the Φ-integral ordering. If $\Phi \subseteq \mathbf{R}^S$, we define *the Φ-integral ordering* on $\Pr(S, \mathcal{A})$, denoted $\preceq_\Phi$, as follows $\mu \preceq_\Phi \nu$ if and only if $\int^* \phi \, d\mu \leq \int^* \phi \, d\nu$ for all $\phi \in \Phi$, where $\int^* f\,d\mu$ denotes the upper μ-integral of f. Then $(\mathrm{Pr}(S,\mathcal{A}),\preceq_\Phi)$ is a proset. If $(\Omega,\mathcal{F},P)$ is a probability space and $X,Y,Z:(\Omega,\mathcal{F})\to(S,\mathcal{A})$ are measurable functions, we let $P_Z(A)=P(Z\in A)$ for $A\in\mathcal{A}$ denote *the distribution* of Z and we write $X\preceq_\Phi Y$ if $P_X\preceq_\Phi P_Y$. Note that $\mu\preceq_\Phi\nu\;\Rightarrow\mu\leq_\Phi\nu$ and that the converse implication holds if $\Phi\subseteq L^1(\mu)\cap L^1(\nu)$. In Section 3 we shall take a closer look at integral orderings,

The classical *stochastic ordering* on $\mathbf{R}$, usually denoted $\preceq_{\mathrm{st}}$, is the integral ordering induced by the indicator functions $\{1_{[a,\infty)}\mid a\in\mathbf{R}\}$; that is $\mu\preceq_{\mathrm{st}}\nu$ if and only if $\mu([a,\infty))\leq\nu([a,\infty))$ for all $a\in\mathbf{R}$. More generally, let $(S,\leq)$ be a proset. Then we let $\mathrm{In}(S,\leq)$ denote the set of all increasing functions from S into $\mathbf{R}$ and we say that $A\subseteq S$ is an *upper interval* if $1_A\in\mathrm{In}(S,\leq)$. We define *the stochastic ordering on* S, denoted $\preceq_{\mathrm{st}}$, to be the integral ordering induced by indicators of upper intervals; that is, $\mu\preceq_{\mathrm{st}}\nu$ if and only if $\mu^*(A)\leq\nu^*(A)$ for every upper interval $A\subseteq S$. If $u\in S$, we define the upper and lower intervals $[u,*]:=\{s\in S\mid s\geq u\}$ and $[*,u]:=\{s\in S\mid s\leq u\}$ and we define *the orthant ordering*, denoted $\preceq_{\mathrm{or}}$, to be the integral ordering induced by $\{1_{[u,*]}\mid u\in S\}$; that is, $\mu\preceq_{\mathrm{or}}\nu$ if and only if $\mu^*([u,*])\leq\nu^*([u,*])$ for all $u\in S$.

Let $k\geq 1$ be an integer. Then we let $\leq$ denote *the product ordering* on $\mathbf{R}^k$; that is, $(x_1,\ldots,x_k)\leq(y_1,\ldots,y_k)$ if and only if $x_i\leq y_i$ for all $i=1,\ldots,k$. If $x=(x_1,\ldots,x_k)$ and $y=(y_1,\ldots,y_k)$ are vectors, we define the lattice infimum and supremum as usual $x\wedge y:=(\min(x_1,y_1),\ldots,\min(x_k,y_k))$ and $x\vee y:=(\max(x_1,y_1),\ldots,\max(x_k,y_k))$, and we define $[x,y]=\{z\in\mathbf{R}^k\mid x\leq z\leq y\}$. We let $\mathcal{B}^k$ denote *the Borel σ-algebra* on $\mathbf{R}^k$. Let $f:\mathbf{R}^k\to\mathbf{R}$ be a given function. Then we say that f is *increasing (decreasing)* if f is increasing (decreasing) with respect to the product ordering $\leq$. We say that f is *supermodular* if $f(x)+f(y)\leq f(x\vee y)+f(x\wedge y)$ for all $x,y\in\mathbf{R}^k$, we say that f is *submodular* if $(-f)$ is supermodular, and we say that f is *modular* if f is supermodular and submodular. We define the following function spaces

$$\mathrm{sm}(\mathbf{R}^k)=\{f\in M(\mathbf{R}^k,\mathcal{B}^k)\mid f \text{ is supermodular }\}$$
$$\mathrm{m}(\mathbf{R}^k)=\{f\in M(\mathbf{R}^k,\mathcal{B}^k)\mid f \text{ is modular }\}\ ,\quad \mathrm{bm}(\mathbf{R}^k)=B(\mathbf{R}^k)\cap\mathrm{m}(\mathbf{R}^k)$$
$$\mathrm{bsm}(\mathbf{R}^k)=B(\mathbf{R}^k)\cap\mathrm{sm}(\mathbf{R}^k)\ ,\quad \mathrm{ism}(\mathbf{R}^k)=\mathrm{In}(\mathbf{R}^k,\leq)\cap\mathrm{sm}(\mathbf{R}^k)$$

and we let $\preceq_{\mathrm{sm}}$, $\preceq_{\mathrm{bm}}$, $\preceq_{\mathrm{m}}$, $\preceq_{\mathrm{bsm}}$ and $\preceq_{\mathrm{ism}}$ denote the integral orderings induced by $\mathrm{sm}(\mathbf{R}^k)$, $\mathrm{bm}(\mathbf{R}^k)$, $\mathrm{m}(\mathbf{R}^k)$, $\mathrm{bsm}(\mathbf{R}^k)$ and $\mathrm{ism}(\mathbf{R}^k)$, respectively. If $k=1$, then every function is supermodular and every increasing function is Borel measurable. Hence, in all dimensions there exists non-measurable supermodular functions and if $k\geq 2$, there exists non-measurable increasing functions. However, in Prop.4.3 below we shall see that an increasing supermodular function is Borel measurable.

Let μ be a Borel probability measure on $\mathbf{R}^k$ and let $F_1,\ldots,F_k:\mathbf{R}\to[0,1]$ denote the one-dimensional marginal distribution functions of μ. Then

$$\overline{F}(x_1,\ldots,x_k):=\min(F_1(x_1),\ldots F_k(x_k))$$

is a k-dimensional distribution function, and if $\lambda_{\overline{F}}$ is the associated Lebesgue-Stieltjes measure, then $\lambda_{\overline{F}}$ is a Borel probability measure on $\mathbf{R}^k$ with the same one-dimensional marginals as μ. By a theorem of A.H. Tchen (see [19]), we have $\int_{\mathbf{R}^k} f\,d\mu \leq \int_{\mathbf{R}^k} f\,d\lambda_{\overline{F}}$ for every supermodular function which is continuous, and satisfies a certain (uniform) integrability condition. In Theorem 4.7, we shall see that $\mu \preceq_{\text{bsm}} \lambda_{\overline{F}}$. In the modern literature it is frequently claimed that $\mu \leq_{\text{sm}} \lambda_{\overline{F}}$ and that $\leq_{\text{sm}}$ coincide with $\leq_{\text{bsm}}$; see for instance [10]. The following example shows that both claims fail when $k \geq 3$.

Example C: (see [17]). Let U be a strictly positive random variable with a one-sided Cauchy distribution; that is, with distribution function F given by:

$$F(x) = \tfrac{2}{\pi}\arctan(x) \text{ if } x > 0 \quad \text{and} \quad F(x) = 0 \text{ if } x \leq 0$$

Since U is strictly positive, we may define $V := \frac{1}{U}$ and $W := \frac{1}{2}|U - V|$. A straightforward computation shows that U, V and W all have distribution function F and so we have $F_U(x) = F_V(x) = F_W(x) = F(x)$ and $F_{(U,U,U)}(x,y,z) = \min(F(x), F(y), F(z))$ for all $x, y, z \in \mathbf{R}$. By [19] and Theorem 4.7 below, we have that $(U,V,W) \preceq_{\text{bsm}} (U,U,U)$ and $(U,V,W) \preceq_{\text{ism}} (U,U,U)$. Set $f(x,y,z) = x + y - 2z$. Then f is continuous, linear and modular and we have

$$f(U,U,U) = 0\ , \qquad f(U,V,W) = 2U1_{\{U<1\}} + \tfrac{2}{U}1_{\{U\geq 1\}}$$
$$0 < f(U,V,W) \leq 2\ , \quad Ef(U,U,U) = 0 < Ef(U,V,W) = \tfrac{2\log 2}{\pi}.$$

Hence, we see that $(U,V,W) \not\preceq_{\text{sm}} (U,U,U)$ and $(U,V,W) \not\preceq_{\text{m}} (U,U,U)$ which shows the integrability condition in Theorem 5 of [19] *cannot* be removed and that $\mu \preceq_{\text{bsm}} \nu$ does *not* imply $\mu \leq_{\text{m}} \nu$.

Let $X = (X_1, \ldots, X_k)$ and $Y = (Y_1, \ldots, Y_k)$ be k-dimensional Gaussian vectors with zero means and covariances $\{\sigma_{ij}^X\}$ and $\{\sigma_{ij}^Y\}$ such that $\sigma_{ii}^Y = \sigma_{ii}^X$ for all $1 \leq i \leq k$ and $\sigma_{ij}^Y \leq \sigma_{ij}^X$ for all $1 \leq i \neq j \leq k$. Let f be a supermodular, locally Lebesgue integrable function. Then we have $\frac{\partial^2 f}{\partial x_i \partial x_j} \geq 0$ in distribution sense for all $1 \leq i \neq j \leq k$. So it is tempting to infer that Slepian's inequality implies $Y \preceq_{\text{sm}} X$. However Slepian's inequality only shows that $Ef(Y) \leq Ef(Y)$ if f satisfies some additional regularity conditions. It can be shown that $Y \preceq_{\text{bsm}} X$; see Theorem 2.8 and Theorem 4.7, but Example C shows that $Y \preceq_{\text{bsm}} X$ does *not* imply $Y \preceq_{\text{sm}} X$ in general, and I don't know if we really have $Y \preceq_{\text{sm}} X$ if X and Y are Gaussian vectors satisfying the above hypotheses. However, Theorem 2.8 and Theorem 4.8 shows that $Ef(Y) \leq Ef(X)$ for a large classs of unbounded supermodular functions. Section 4 is devoted the study of the modular orderings introduced above.

2. Slepian's inequality

In this section I shall prove a general version of Slepian's inequality where the partial derivatives are understood in the sense of Schwartz distributions. The idea is to approximate the function $f : \mathbf{R}^k \to \mathbf{R}$ with infinitely often differentiable functions $f_1, f_2, \dots$ satisfying Slepian's inequality. The approximating sequence will taken as the convolution integrals $f_n(x) = \int_K f(x + \frac{y}{n}) g(y)\, dy$ where $K \subseteq \mathbf{R}^k$ is a compact starshaped set and g is a nonnegative infinitely often differentiable function satisfying $\{g \neq 0\} \subseteq K$ and $\int_K g(y)\, dy = 1$. Below we shall see that if f is locally Lebesgue integrable, then f_n is an infinitely often differentiable function inheriting many properties of f and that $f_n(x) \to f(x)$ for all x in a large subset of $\mathbf{R}^k$. However, this requires some preparatory definitions and lemmas.

Let S be a set and let $\kappa : S \to [0, \infty]$ be a given function. If $f \in \mathbf{R}^S$, we let $||f||_\kappa := \inf\{c \in \mathbf{R}_+ \mid |f(s)| \leq c\, \kappa(s)\ \forall\, s \in S\}$ denote *the weighted sup-norm* of $f \in \mathbf{R}^S$ with the usual convention $\inf \emptyset := \infty$. If $\Phi \subseteq \mathbf{R}^S$ is a set of functions, we let $\Phi_+ := \Phi \cap \mathbf{R}_+^S$ denote the set of all nonnegative functions in Φ. If S and T are topological spaces and $\phi : S \to T$ is a given function, we let $C(\phi)$ denote *the continuity set* of ϕ; that is, the set of all $s \in S$ such that ϕ is continuous at s.

Let $k \geq 1$ be an integer and set $[k] := \{1, \dots, k\}$. We let $e_1, \dots, e_k$ denote the standard unit vectors in $\mathbf{R}^k$. If $x = (x_1, \dots, x_k) \in \mathbf{R}^k$ and $y = (y_1, \dots, y_k) \in \mathbf{R}^k$, we let $\langle x, y \rangle = \sum_{i=1}^k x_i y_i$ denote *the inner product* and we let $||x|| = \langle x, x \rangle^{1/2}$ denote *the Euclidian norm.* We let λ_k denote the k-dimensional *Lebesgue measure* on $\mathbf{R}^k$. We say that $f : \mathbf{R}^k \to \mathbf{R}$ is *locally bounded* if f is bounded on every compact subset of $\mathbf{R}^k$, we say that f is *locally λ_k-integrable* if $1_C f \in L^1(\lambda_k)$ for every compact set $C \subseteq \mathbf{R}^k$, and we let $L^1_{\mathrm{loc}}(\lambda_k)$ denote the set of all locally λ_k-integrable functions.

Let $f : \mathbf{R}^k \to \mathbf{R}$ be a given function. If $i \in [k]$ and $t \in \mathbf{R}$, we let $\Delta_i^t f(x) := f(x + te_i) - f(x)$ for $x \in \mathbf{R}^k$ denote the usual *difference operator.* If $\theta \in \mathbf{R}^k$, we say that f is *θ-differentiable at x* if $t \curvearrowright f(x + t\theta)$ is differentiable at 0 and if so we let $\frac{\partial f}{\partial \theta}(x) := \lim_{t \to 0} t^{-1} (f(x + t\theta) - f(x))$ denote *the directional θ-derivative* of f at x. In particular, we let $\frac{\partial f}{\partial x_i}(x) := \lim_{t \to 0} t^{-1}\, \Delta_i^t f(x)$ denote the partial derivative whenever it exists. We say that f is *partially differentiable at x* if the partial derivatives $\frac{\partial f}{\partial x_i}(x)$ exists for all $i \in [k]$ and if so we let $\nabla f(x) := (\frac{\partial f}{\partial x_1}(x), \dots, \frac{\partial f}{\partial x_k}(x))$ denote the *gradient of f*. We say that f is *θ-differentiable* if f is θ-differentiable at all $x \in \mathbf{R}^k$ and we say that f is *continuously θ-differentiable* if f is θ-differentiable and $x \curvearrowright \frac{\partial f}{\partial \theta}(x)$ is continuous on $\mathbf{R}^k$. We say that f is *partially differentiable* if f is partially differentiable at all $x \in \mathbf{R}^k$. Recall that f is *Fréchet differentiable at x* if the directional derivative $\frac{\partial f}{\partial \theta}(x)$ exists for all $\theta \in \mathbf{R}^k$ and $\frac{\partial f}{\partial \theta}(x) = \langle \theta, \nabla f(x) \rangle$ for all $\theta \in \mathbf{R}^k$. Recall that f is *differentiable at x with differential $D \in \mathbf{R}^k$* if $\lim_{\theta \to 0} \|\theta\|^{-1} |f(x + \theta) - f(x) - \langle D, \theta \rangle| = 0$.

If $i_1, \dots, i_p \in [k]$, we let $\frac{\partial^p f}{\partial x_{i_1} \cdots \partial x_{i_p}}(x)$ denote the pth order partial derivative whenever it exists. We let $C^\infty(\mathbf{R}^k)$ denote the set of all infinitely often differentiable functions $f : \mathbf{R}^k \to \mathbf{R}$ and if $\kappa : \mathbf{R}^k \to [0, \infty]$ is a nonnegative function we let $C_\kappa^\infty(\mathbf{R}^k)$ denote the set of all $f \in C^\infty(\mathbf{R}^k)$ satisfying

$$\|f\|_\kappa < \infty \quad \text{and} \quad \left\| \frac{\partial^p f}{\partial x_{i_1} \cdots \partial x_{i_p}} \right\|_\kappa < \infty \ \ \forall p \geq 1\ \forall i_1, \dots, i_p \in [k].$$

In particular, we let $C_b^\infty(\mathbf{R}^k)$ denote the set of all bounded, infinitely often differentiable functions with bounded derivatives of all orders. We let $C_{\circ\circ}^\infty(\mathbf{R}^k)$ denote the set of all $f \in C^\infty(\mathbf{R}^k)$ with compact support and we let ϖ denote the usual inductive limit topology on $C_{\circ\circ}^\infty(\mathbf{R}^k)$; see [13]. We let $\mathcal{D}(\mathbf{R}^k)$ denote the set all *Schwartz distributions*; that is, the set of all ϖ-continuous linear functionals $\zeta : C_{\circ\circ}^\infty(\mathbf{R}^k) \to \mathbf{R}$. If $\zeta \in \mathcal{D}(\mathbf{R}^k)$, we write $\zeta \geq 0$ if and only if $\zeta(\phi) \geq 0$ for all $\phi \in C_{\circ\circ}^\infty(\mathbf{R}^k)_+$. If $f \in L^1_{\mathrm{loc}}(\lambda_k)$, then $f(\phi) := \int_{\mathbf{R}^k} f(x)\,\phi(x)\,dx$ for $\phi \in C_{\circ\circ}^\infty(\mathbf{R}^k)$ defines a Schwartz distribution corresponding to f and if $i_1, \dots, i_p \in [k]$, then

$$\partial_{i_1,\dots,i_p} f(\phi) := (-1)^p \int_{\mathbf{R}^k} f(x)\, \frac{\partial^p \phi}{\partial x_{i_1} \cdots \partial x_{i_p}}(x)\,dx \text{ for } \phi \in C_{\circ\circ}^\infty(\mathbf{R}^k)$$

defines a Schwartz distributions, which corresponds to the "the partial derivative" $\frac{\partial^p f}{\partial x_{i_1} \cdots \partial x_{i_p}}$.

Recall that $K \subseteq \mathbf{R}^k$ is *starshaped* if $0 \in K$ and $\alpha x \in K$ for all $x \in K$ and all $0 \leq \alpha \leq 1$. Let $K \subseteq \mathbf{R}^k$ be a bounded, starshaped Borel set and let $x \in \mathbf{R}^k$ be a given vector. Then we say that f is *continuous at x along K* if

$$\lim_{n\to\infty} \Big\{ \sup_{y \in K} |f(x + \tfrac{y}{n}) - f(x)| \Big\} = 0 \tag{2.1}$$

and we let $C^K(f)$ denote the set of all $x \in \mathbf{R}^k$ satisfying (2.1). If 0 belongs to the interior of K, then continuity along K coincides with ordinary continuity. We say that f is *right continuous at x* if f is continuous at x along the unit cube $[0,1]^k$, say that f is *left continuous at x* if f is continuous at x along the negative unit cube $[-1,0]^k$.

Let $K \subseteq \mathbf{R}^k$ be a bounded, starshaped Borel set. Then we say that f is *approximately continuous at x along K* if f is locally λ_k-integrable and

$$\lim_{n\to\infty} \int_K |f(x + \tfrac{y}{n}) - f(x)|\,dy = 0. \tag{2.2}$$

We let $C^K_{\mathrm{ap}}(f)$ denote the set of all $x \in \mathbf{R}^k$ satisfying (2.2). Let $f \in L^1_{\mathrm{loc}}(\lambda_k)$ be a Borel function. By the Fubini-Tonelli theorem, we see that $C^K_{\mathrm{ap}}(f)$ is a Borel set containing $C^K(f)$ and by Theorem III.12.8 p. 217 in [1] we have $\lambda_k(\mathbf{R}^k \setminus C^K_{\mathrm{ap}}(f)) = 0$.

If $f, g : \mathbf{R}^k \to \mathbf{R}$ are λ_k-measurable and $\int_{\mathbf{R}^k} |f(x-y)\,g(y)|\,dy < \infty$ for all $x \in \mathbf{R}^k$, we say that *the convolution* exists and we let $(f \star g)(x) := \int_{\mathbf{R}^k} f(x-y)\,g(y)\,dy$ denote *the convolution* of f and g.

Lemma 2.1. *Let $f : \mathbf{R}^k \to \mathbf{R}$ be a locally λ_k-integrable function and let $g : \mathbf{R}^k \to \mathbf{R}$ be a bounded Lebesgue measurable function with compact support. Then the convolution $h(x) := (f \star g)(x)$ exists and is continuous on $\mathbf{R}^k$ and if $\theta \in \mathbf{R}^k$ is a given vector, we have*

(1) *If f is θ-differentiable and $\frac{\partial f}{\partial \theta} \in L^1_{\mathrm{loc}}(\lambda_k)$, then h is continuously θ-differentiable and we have $\frac{\partial h}{\partial \theta}(x) = (\frac{\partial f}{\partial \theta} \star g)(x) \quad \forall x \in \mathbf{R}^k$.*

(2) *If g is θ-differentiable and $\frac{\partial g}{\partial\theta}$ is bounded, then h is continuously θ-differentiable and we have $\frac{\partial h}{\partial\theta}(x) = (f \star \frac{\partial g}{\partial\theta})(x) \ \ \forall x \in \mathbf{R}^k$.*

(3) *If f and g are θ-differentiable, $\frac{\partial f}{\partial\theta} \in L^1_{\text{loc}}(\lambda_k)$ and $\frac{\partial g}{\partial\theta}$ is bounded, then we have*

$$\int_{\mathbf{R}^k} \tfrac{\partial f}{\partial\theta}(x) \cdot g(y)\, dy = -\int_{\mathbf{R}^k} f(y) \cdot \tfrac{\partial g}{\partial\theta} g(y)\, dy.$$

Proof. Set $B_r := \{x \in \mathbf{R}^k \mid ||x|| \le r\}$ for $r \ge 0$. Since g is bounded with compact support, there exist $a, \rho > 0$ such that $|g(x)| \le a$ for all $x \in \mathbf{R}^k$ and $g(x) = 0$ for all $x \notin B_\rho$. Since f is locally λ_k-integrable and $|f(x-y)\, g(y)| \le a\, |f(x-y)|\, 1_{B_\rho}(y)$, we see that the convolution $h(x) = (f \star g)(x)$ exists for all $x \in \mathbf{R}^k$. Let $r > 0$ and $x \in B_r$ be given. Then $f_r := f\, 1_{B_{r+\rho}} \in L^1(\lambda_k)$ and we have $f_r(x-y)\, g(y) = f(x-y)\, g(y)$ for all $y \in \mathbf{R}^k$. Hence, we have $(f_r \star g)(x) = h(x)$ for all $x \in B_r$ and by Theorem 1.1.6 p. 4 in [14], we have that $f_r \star g$ is continuous on $\mathbf{R}^k$. Hence, we see that h is continuous on $\mathbf{R}^k$.

Suppose that f is θ-differentiable and that $\frac{\partial f}{\partial\theta} \in L^1_{\text{loc}}(\lambda_k)$. Let $x \in \mathbf{R}^k$ be given. By the argument above we have that the convolutions $\frac{\partial f}{\partial\theta} \star g$ and $|\frac{\partial f}{\partial\theta}| \star 1_{B_r}$ exist and are continuous on $\mathbf{R}^k$ for all $r \ge 0$. Let $x \in \mathbf{R}^k$ be given. By the Fubini-Tonelli theorem and locally boundedness of $|\frac{\partial f}{\partial\theta}| \star 1_{B_r}$, there exists a λ_k-null set N_x such that $s \curvearrowright \frac{\partial f}{\partial\theta}(x - y + s\theta)$ is locally λ_1-integrable on $\mathbf{R}$ for all $y \notin N_x$ and we have

$$\begin{aligned}\int_0^t (\tfrac{\partial f}{\partial\theta} \star g)(x + s\theta)\, ds &= \int_0^t ds \int_{\mathbf{R}^k} \tfrac{\partial f}{\partial\theta}(x - y + s\theta) g(y)\, dy \\ &= \int_{\mathbf{R}^k} dy \int_0^t \tfrac{\partial f}{\partial\theta}(x - y + s\theta)\, g(y)\, ds.\end{aligned}$$

Let $y \in \mathbf{R}^k \setminus N_x$ be given and set $F_{x,y}(s) := f(x-y+s\theta)$. Then $F_{x,y}$ is differentiable with derivative $F'_{x,y}(s) = \frac{\partial f}{\partial\theta}(x - y + s\theta)$ and $F'_{x,y}$ is locally λ_1-integrable. By a classical theorem of Denjoy and Banach (see Thm. IX.4.5 p. 271 and Thm. IX.7.4 p. 284 in [15]), we see that $F_{x,y}$ is absolutely continuous with Lebesgue derivative $F'_{x,y}$. In particular, we have $F_{x,y}(t) - F_{x,y}(0) = \int_0^t F'_{x,y}(s)\, ds$ and since $\int_{\mathbf{R}^k} F_{x,y}(s)\, g(y)\, dy = h(x + s\theta)$, we have

$$\int_0^t (\tfrac{\partial f}{\partial\theta} \star g)(x + s\theta)\, ds = \int_{\mathbf{R}^k} (F_{x,y}(t) - F_{x,y}(0))\, g(y)\, dy = h(x + t\theta) - h(x).$$

Since $\frac{\partial f}{\partial\theta} \star g$ is continuous, we see that h is continuously θ-differentiable with $\frac{\partial h}{\partial\theta}(x) = (\frac{\partial f}{\partial\theta} \star g)(x)$ for all $x \in \mathbf{R}^k$. Thus, (1) is proved and (2) follows in the same manner. Applying (1) and (2) on $f(y)$ and $g(-y)$ with $x = 0$, we obtain (3). $\square$

Lemma 2.2. *Let $f : \mathbf{R}^k \to \mathbf{R}$ be a locally λ_k-integrable function and let $K \subseteq \mathbf{R}^k$ be starshaped, bounded Borel set. Let $g \in C^\infty_+(\mathbf{R}^k)$ be given such that $\{g \ne 0\} \subseteq K$ and $\int_K g(y)\, dy = 1$. Set $g_n(x) := n^k\, g(-nx)$ and $f_n(x) := (f \star g_n)(x)$ for $n \ge 1$*

and $x \in \mathbf{R}^k$; see Lemma 2.1. Let $p \geq 1$ and $i_1, \ldots, i_p \in [k]$ be given integers and let us define

$$\kappa(x) = \sup_{y \in K} |f(x+y)| \quad \forall x \in \mathbf{R} \ , \ c_{i_1 \ldots i_p} = \sup_{y \in K} |\tfrac{\partial^p g}{\partial x_{i_1} \cdots \partial x_{i_p}}(y)|.$$

Then $0 \leq c_{1_1,\ldots,i_p} < \infty$ and we have

(1) $f_n \in C^\infty(\mathbf{R}^k)$ *and* $f_n(x) = \int_{\mathbf{R}^k} f(x + \frac{y}{n}) g(y) dy \quad \forall x \in \mathbf{R}^k \ \forall n \geq 1.$

(2) $\frac{\partial^p f_n}{\partial x_{i_1} \cdots \partial x_{i_p}}(x) = (-n)^p \int_{\mathbf{R}^k} f(x + \frac{y}{n}) \frac{\partial^p g}{\partial x_{i_1} \cdots \partial x_{i_p}}(y) \, dy \quad \forall x \in \mathbf{R}^k \ \forall n \geq 1.$

(3) $|f_n(x)| \leq \kappa(x)$ *and* $|\frac{\partial^p f_n}{\partial x_{i_1} \cdots \partial x_{i_p}}(x)| \leq c_{i_1 \ldots i_p} n^p \kappa(x) \quad \forall x \in \mathbf{R}^k \ \forall n \geq 1.$

(4) $\lim_{n \to \infty} f_n(x) = f(x) \quad \forall x \in C_{\mathrm{ap}}^K(f).$

(5) $\lim_{n \to \infty} \left\{ \sup_{y \in C} \int_C |f(x + \frac{y}{n}) - f(x)| dx \right\} = 0$ *for all compact sets* $C \subseteq \mathbf{R}^k$.

(6) *If f is bounded with compact support, we have*

$$\lim_{n \to \infty} \int_{\mathbf{R}^k} |(f_n(x) - f(x)) \psi(x)| dx = 0 \quad \forall \psi \in L^1_{\mathrm{loc}}(\lambda_k)$$

.

Proof. (1)–(2): Note that g, g_n, $G = \frac{\partial^p g}{\partial x_{i_1} \cdots \partial x_{i_p}}$ and $G_n = \frac{\partial^p g_n}{\partial x_{i_1} \cdots \partial x_{i_p}}$ are infinitely often differentiable with compact supports and we have

$$G_n(x) = (-1)^p \, n^{k+p} \, G(-nx).$$

So by Lemma 2.1 we see that $f_n \in C^\infty(\mathbf{R}^k)$ and that $\frac{\partial^p f_n}{\partial x_{i_1} \cdots \partial x_{i_p}}(x) = f \star G_n$. Hence, we see that (1)–(2) follows from the substitution $z = -\frac{y}{n}$.

(3): Let $n \geq 1$ be given. Since K is starshaped, we have $|f(x + \frac{y}{n})| \leq \kappa(x)$ for all $(x,y) \in \mathbf{R}^k \times K$ and since $g \geq 0$ and $\int_K g(y) \, dy = 1$, we see that (3) follows from (1)–(2).

(4): By (1), we have $|f_n(x) - f(x)| \leq a \int_K |f(x + \frac{y}{n}) - f(x)| \, dy$ where $a := \sup_{y \in K} g(y)$. Since $a < \infty$, we see that (4) holds.

(5): Let $r, \varepsilon > 0$ be given and set $B_r := \{x \in \mathbf{R}^k \mid ||x|| < r\}$. Since $f \in L^1_{\mathrm{loc}}(\lambda_k)$, we have $f_r := 1_{B_{1+r}} f \in L^1(\lambda_k)$. By Theorem 1.1.5 in [14], there exists $0 < \delta < 1$ such that $\int_{\mathbf{R}^k} |f_r(x+u) - f_r(x)| \, dx < \varepsilon$ for all $||u|| \leq \delta$. Let $x, y \in B_r$ and $n \geq \frac{r}{\delta}$ be given. Since $||\frac{y}{n}|| \leq \delta < 1$, we have $x \in B_{r+1}$ and $x + \frac{y}{n} \in B_{r+1}$ and so we have $f(x) = f_r(x)$ and $f(x + \frac{y}{n}) = f_r(x + \frac{y}{n})$. Hence, we have

$$\int_{B_r} |f(x + \tfrac{y}{n}) - f(x)| dx \leq \int_{\mathbf{R}^k} |f_r(x + \tfrac{y}{n}) - f_r(x)| dx \leq \varepsilon$$

for all $n \geq \frac{r}{\delta}$ and all $y \in B_r$. Since $r > 0$ is arbitrary, we see that (5) holds.

(6): Suppose that f is bounded with compact support. Then

$$b := \sup_{x \in \mathbf{R}^k} |f(x)| < \infty$$

and there exists $r > 0$ such that $\{f \neq 0\} \cup \{g \neq 0\} \subseteq B_r$. By (1), we see that $\{f_n \neq 0\} \subseteq B_{2r}$ and that $|f_n(x)| \leq b\, 1_{B_{2r}}(x)$. By (4), we see that $f_n(x) \to f(x)$ λ_k-a.e. and so we see that (6) follows from Lebesgue's convergence theorem. □

Lemma 2.3. *Let $f : \mathbf{R}^k \to \mathbf{R}$ be a locally λ_k-integrable function and let $K \subseteq \mathbf{R}^k$ be a starshaped, bounded Borel set. Let $g \in C_+^\infty(\mathbf{R}^k)$ be given such that $\{g \neq 0\} \subseteq K$ and $\int_K g(y)\,dy = 1$ and set $f_n(x) = \int_K f(x + \frac{y}{n})\, g(y)\,dy$ for $n \geq 1$ and $x \in \mathbf{R}^k$; see Lemma 2.2*

Let $A = \{a_{ij}\}$ be a $(k \times k)$-matrix and let $\mathcal{H}^A$ denote the set of all twice partially differentiable functions $h : \mathbf{R}^k \to [0, \infty)$ with compact support such that h, $\frac{\partial h}{\partial x_i}$ and $\frac{\partial^2 h}{\partial x_i \partial x_j}$ are locally λ_k-integrable for all $i, j \in [k]$ and $\sum_{i=1}^k \sum_{j=1}^k a_{ij} \frac{\partial^2 h}{\partial x_i \partial x_j}$ is bounded. Let $\epsilon_1, \delta_1, \epsilon_2, \delta_2, \ldots > 0$ be positive numbers such that $\epsilon_n \to 0$ and $\delta_n \to 0$. Then the following four statements are equivalent:

(1) $\sum_{i=1}^k \sum_{j=1}^k a_{ij} \partial_{ij} f \geq 0.$

(2) $\sum_{i=1}^k \sum_{j=1}^k a_{ij} \int_{\mathbf{R}^k} f(x) \frac{\partial^2 h}{\partial x_i \partial x_j}(x) dx \geq 0 \quad \forall h \in \mathcal{H}^A.$

(3) $\sum_{i=1}^k \sum_{j=1}^k a_{ij} \frac{\partial^2 f_n}{\partial x_i \partial x_j}(x) \geq 0 \quad \forall x \in \mathbf{R}^k\ \forall n \geq 1.$

(4) $\sum_{i=1}^k \sum_{j=1}^k a_{ij} \Delta_i^{\epsilon_n} \Delta_j^{\delta_n} f(x) \geq 0 \quad \lambda_k$*-a.e.* $\quad \forall n \geq 1.$

In particular, we have

(5) *If f is convex and A is nonnegative definite, then* $\sum_{i=1}^k \sum_{j=1}^k a_{ij} \partial_{ij} f \geq 0$

and if f is twice partially differentiable and f, $\frac{\partial f}{\partial x_i}$ and $\frac{\partial^2 f}{\partial x_i \partial x_j}$ are locally λ_k-integrable for all $i, j \in [k]$, then we have

(6) $\sum_{i=1}^k \sum_{j=1}^k a_{ij} \partial_{ij} f \geq 0 \ \Leftrightarrow\ \sum_{i=1}^k \sum_{j=1}^k a_{ij} \frac{\partial^2 f}{\partial x_i \partial x_j}(x) \geq 0 \quad \lambda_k$*-a.e.*

Proof. Set $\zeta = \sum_{i=1}^k \sum_{j=1}^k a_{ij}\, \partial_{ij} f$, $F_n(x) = \sum_{i=1}^k \sum_{j=1}^k a_{ij} \frac{\partial^2 f_n}{\partial x_i \partial x_j}(x)$ and $g_n(x) = n^k g(-nx)$ for all $x \in \mathbf{R}^k$ and all $n \geq 1$.

(1) $\Rightarrow$ (2): Suppose that $\zeta \geq 0$ and let $h \in \mathcal{H}^A$ be given. Set $H(x) = \sum_{i=1}^k \sum_{j=1}^k a_{ij} \frac{\partial^2 h}{\partial x_i \partial x_j}(x)$. By Lemma 2.2 and local λ_k-integrability of h and H, we have that the convolutions $h_n := h \star g_n$ and $H_n := H \star g_n$ exist and satisfies (1)–(6) in Lemma 2.2. Since h is nonnegative with compact support, we have $h_n, H_n \in$

$C_{\circ\circ}^\infty(\mathbf{R}^k)$ and $h_n \geq 0$ and by Lemma 2.1, we have $H_n = \sum_{i,j} a_{ij} \frac{\partial^2 h_n}{\partial x_i \partial x_j}$. Since $\zeta \geq 0$, we have $0 \leq \zeta(h_n) = \int_{\mathbf{R}^k} f H_n \, d\lambda_k$ for all $n \geq 1$. Recall that H is bounded with compact support and that f is locally λ_k-integrable. So by Lemma 2.2.(6) applied to the pair $(f, \psi) := (H, f)$, we have $\int_{\mathbf{R}^k} f H \, d\lambda_k = \lim_{n\to\infty} \zeta(h_n) \geq 0$.

(2) $\Rightarrow$ (3): Suppose that (2) holds and let $x \in \mathbf{R}^k$ and $n \geq 1$ be given. Set $g_{nx}(y) = g_n(x-y)$ for $y \in \mathbf{R}^k$. By Lemma 2.1, we have

$$\frac{\partial^2 f_n}{\partial x_i \partial x_j}(x) = \int_{\mathbf{R}^k} f(y) \frac{\partial^2 g_n}{\partial x_i \partial x_j}(x-y) dy = \int_{\mathbf{R}^k} f(y) \frac{\partial^2 g_{nx}}{\partial x_i \partial x_j}(y) dy$$

and since $g_{nx} \in \mathcal{H}^A$, we see that (2) implies (3).

(3) $\Rightarrow$ (4): Suppose that (3) holds and let $n \geq 1$ and $u, v > 0$ be given. Since $f_n \in C^\infty(\mathbf{R}^k)$, we have

$$\Delta_i^u \Delta_j^v f_n(x) = \int_0^u ds \int_0^v \frac{\partial^2 f_n}{\partial x_i \partial x_j}(x + se_i + te_j) \, dt \quad \forall x \in \mathbf{R}^k.$$

So by (3) we have $\sum_{i,j} a_{ij} \, \Delta_i^u \Delta_j^v f_n(x) \geq 0$ for all $x \in \mathbf{R}^k$ and by Lemma 2.2.(4) we have that $f_n \to f$ λ_k-a.e. Hence, we see that (3) implies (4).

(4) $\Rightarrow$ (1): Suppose that (4) holds. As above, we see that (4) implies $F_n(x) \geq 0$ for all $x \in \mathbf{R}^k$. Let $h \in C_{\circ\circ}^\infty(\mathbf{R}^k)_+$ be given and set $H = \sum_{i,j} a_{ij} \frac{\partial^2 h}{\partial x_i \partial x_j}$. Then we have $\zeta(h) = \int_{\mathbf{R}^k} f H \, d\lambda_k$ and observe that H is bounded with compact support. So by Lemma 2.1.(3) and Lemma 2.2.(6) we have

$$\zeta(h) = \lim_{n\to\infty} \int_{\mathbf{R}^k} f_n(x) H(x) \, dx = \lim_{n\to\infty} \int_{\mathbf{R}^k} F_n(x) h(x) \, dx \geq 0.$$

Hence we see that (4) implies (1).

(5): Suppose that f is convex and A is nonnegative definite. Let $n \geq 1$ be a given integer. By nonnegativity of g, we see that f_n is convex and infinitely often differentiable. Let $x \in \mathbf{R}^k$ and $n \geq 1$ be given and set $b_{ij} = \frac{\partial^2 f_n}{\partial x_i \partial x_j}(x)$ for $i, j \in [k]$. Then $B = (b_{ij})$ is the Hessian of f_n and since f_n is convex, we have that B is a nonnegative definite $(k \times k)$-matrix. By Schur's product theorem (see Thm. 7.5.3 p. 458 in [5]) we have that the Hadamard product $(c_{ij}) = (a_{ij} \, b_{ij})$ is nonnegative definite. In particular, we have

$$\sum_{i=1}^k \sum_{j=1}^k a_{ij} \frac{\partial^2 f_n}{\partial x_i \partial x_j}(x) = \sum_{i=1}^k \sum_{j=1}^k c_{ij} \geq 0.$$

Hence, we see that (5) follows from the equivalence of (3) and (1).

(6): Suppose that f is twice partially differentiable such that $\frac{\partial f}{\partial x_i}$ and $\frac{\partial^2 f}{\partial x_i \partial x_j}$ are locally λ_k-integrable for all $i, j \in [k]$. Then $F := \sum_{i,j} a_{ij} \frac{\partial^2 f}{\partial x_i \partial x_j}$ belongs to $L^1_{\mathrm{loc}}(\lambda_k)$ and by Lemma 2.1 we have

$$\zeta(\phi) = \sum_{i=1}^k \sum_{j=1}^k a_{ij} \int_{\mathbf{R}^k} f(x) \frac{\partial^2 \phi}{\partial x_i \partial x_j}(x) dx = \int_{\mathbf{R}^k} F(x) \phi(x) \, dx \ \forall \phi \in C_{\circ\circ}^\infty(\mathbf{R}^k).$$

Since $F \geq 0$ λ_k-a.e. if and only if $\int_{\mathbf{R}^k} F\phi \, d\lambda_k \geq 0$ for all $\phi \in C_{\circ\circ}^{\infty}(\mathbf{R}^k)_+$, we see that (6) holds. □

Lemma 2.4. *Let $f : \mathbf{R}^k \to \mathbf{R}$ be a locally λ_k-integrable Borel function and let $\{a_{ij}\}_{1\leq i,j\leq k}$ be a $(k \times k)$-matrix. Let $\theta = (\theta_1, \ldots, \theta_k)$, $b = (b_1, \ldots, b_k)$ and $c = (c_1, \ldots, c_k)$ be a given vectors. If there exist functions $h : \mathbf{R}^k \to \mathbf{R}$ and $\psi : \mathbf{R} \to \mathbf{R}$ satisfying $f(x + t\theta) = h(x) + \psi(t)$ for all $x \in \mathbf{R}^k$ and all $t \in \mathbf{R}$, then we have*

(1) $f(x + t\theta) = f(x) + \gamma t \quad \forall x \in \mathbf{R}^k \ \forall t \in \mathbf{R} \quad$ *where* $\gamma := f(\theta) - f(0)$.

(2) $\sum_{i=1}^{k} \sum_{j=1}^{k} a_{ij}\partial_{ij} f = \sum_{i=1}^{k} \sum_{j=1}^{k} (a_{ij} + b_i\theta_j + c_j\theta_i)\partial_{ij} f.$

Proof. (1): Set $\psi_0(t) = \psi(t) - \psi(0)$ for all $t \in \mathbf{R}$. Since $f(x) = h(x) + \psi(0)$, we have $f(x + t\theta) = f(x) + \psi_0(t)$ and so we have

$$\psi_0(s + t) = f(s\theta + t\theta) - f(0) = f(s\theta) + \psi_0(t) - f(0) = \psi_0(s) + \psi_0(t)$$

for all $s, t \in \mathbf{R}$. Since f is Borel measurable and $\psi_0(t) = f(t\theta) - f(0)$, we see that ψ_0 is a Borel function satisfying $\psi_0(0) = 0$ and $\psi_0(s + t) = \psi_0(s) + \psi_0(t)$ for all $s, t \in \mathbf{R}$. So by [11] we have $\psi_0(t) = \gamma t$ for all $t \in \mathbf{R}$ where $\gamma = \psi_0(1) = f(\theta) - f(0)$. Since $f(x + t\theta) = f(x) + \psi_0(t)$, we see that (1) holds.

(2): Let $g \in C_{\circ\circ}^{\infty}(\mathbf{R}^k)_+$ be a nonnegative function with $\int_{\mathbf{R}^k} g(y)\, dy = 1$ and set $f_n(x) = \int_{\mathbf{R}^k} f(x + \frac{y}{n})\, g(y)\, dy$ for all $x \in \mathbf{R}^k$ and all $n \in \mathbf{N}$ (see Lem. 2.2). Then $f_n \in C^{\infty}(\mathbf{R}^k)$ and by (1), we have $f_n(x + t\theta) = f_n(x) + \gamma t$. In particular, we see that $\gamma = \frac{\partial f_n}{\partial \theta}(x) = \sum_{i=1}^{k} \theta_i \frac{\partial f_n}{\partial x_i}(x)$ for all $x \in \mathbf{R}^k$. Hence, we have

$$\sum_{j=1}^{k} \sum_{i=1}^{k} \theta_i c_j \frac{\partial^2 f_n}{\partial x_j \partial x_i}(x) = \sum_{j=1}^{k} c_j \frac{\partial}{\partial x_j} \left\{ \sum_{i=1}^{k} \theta_i \frac{\partial f_n}{\partial x_i}(x) \right\} = 0 \quad \forall x \in \mathbf{R}^k.$$

In the same manner we see that $\sum_{i=1}^{k} \sum_{j=1}^{k} b_i\theta_j \frac{\partial^2 f_n}{\partial x_i \partial x_j}(x) = 0$ for all $x \in \mathbf{R}^k$ and so we see that (2) follows from Proposition 2.3. □

Lemma 2.5. *Let $\phi, \psi : \mathbf{R} \to \mathbf{R}$ be absolutely continuous functions with Lebesgue derivatives $\dot{\phi}(t)$ and $\dot{\psi}(t)$. Set $\psi_*(t) := \int_t^{\infty} |\dot{\psi}(s)|\, ds$ for $t \geq 0$ and $\psi_*(t) := \int_{-\infty}^{t} |\dot{\psi}(s)|\, ds$ for $t < 0$. If (ϕ, ψ) satisfies the following condition*

(1) $\dot{\psi} \in L^1(\lambda_1)$, $\dot{\phi} \cdot \psi_* \in L^1(\lambda_1)$ *and* $\lim_{x\to\infty} \psi(x) = 0 = \lim_{x\to-\infty} \psi(x)$

then $\dot{\phi} \cdot \psi$ and $\phi \cdot \dot{\psi}$ are λ-integrable and we have

(2) $\int_{-\infty}^{\infty} \dot{\phi}(s)\psi(s)\, ds = -\int_{-\infty}^{\infty} \phi(s)\dot{\psi}(s)\, ds.$

Proof. Since $\dot{\psi}$ is λ_1-integrable and $\lim_{x\to\pm\infty} \psi(x) = 0$, we have

$$\psi(t) = \int_{-\infty}^{t} \dot{\psi}(s)\, ds = -\int_{t}^{\infty} \dot{\psi}(s)\, ds$$

for all $t \in \mathbf{R}$. In particular, we see that $|\psi(t)| \le \psi_*(t)$ for all $t \in \mathbf{R}$ and so by (1) we have $\dot\phi(t)\,\psi(t) \in L^1(\lambda)$. By the Fubini-Tonelli theorem, we have

$$\begin{aligned}\int_0^\infty |\phi(t)-\phi(0)|\cdot|\dot\psi(t)|dt &\le \int_0^\infty dt\int_0^t |\dot\phi(s)\dot\psi(t)|\,ds\\ &= \int_0^\infty ds\int_s^\infty |\dot\phi(s)\dot\psi(t)|dt = \int_0^\infty |\dot\phi(s)|\psi_*(s)\,ds < \infty\end{aligned}$$

and in the same manner, we see that $\int_{-\infty}^0 |\phi(t)-\phi(0)|\cdot|\dot\psi(t)|\,dt < \infty$. Since $\dot\psi \in L^1(\lambda)$, we see that $\phi(t)\,\dot\psi(t) \in L^1(\lambda)$. So by the Fubini-Tonelli theorem we have

$$\begin{aligned}\int_0^\infty \dot\phi(t)\psi(t)\,dt &= -\int_0^\infty dt\int_t^\infty \dot\phi(t)\dot\psi(s)\,ds = -\int_0^\infty ds\int_0^s \dot\phi(t)\dot\psi(s)\,dt\\ &= \int_0^\infty (\phi(0)-\phi(s))\dot\psi(s)\,ds = -\phi(0)\psi(0) - \int_0^\infty \phi(s)\dot\psi(s)\,ds.\end{aligned}$$

In the same manner, we see that $\int_{-\infty}^0 \dot\phi(t)\,\psi(t)\,dt = \phi(0)\psi(0) - \int_{-\infty}^0 \phi(t)\,\dot\psi(t)\,dt$. Adding the two equalities we obtain (2). □

Lemma 2.6. *Let $(U_0, U_1, \ldots, U_k)$ be a $(k+1)$-dimensional Gaussian random vector with mean zero and set $U := (U_1, \ldots, U_k)$ and $\theta := (\theta_1, \ldots, \theta_k)$ where $\theta_i := E(U_0\,U_i)$ for $i = 1, \ldots, k$. Let $h : \mathbf{R}^k \to \mathbf{R}$ be a θ-differentiable Borel function satisfying $E|U_0 h(U)| < \infty$ and $E|\frac{\partial h}{\partial\theta}(U)| < \infty$. Then we have*

(1) $E\{U_0 h(U)\} = E\{\frac{\partial h}{\partial\theta}(U)\}$.

Proof. Set $\sigma^2 = EU_0^2$. If $\sigma^2 = 0$, then we have $U_0 = 0$ a.s. and so we have $\theta = 0$ and $\frac{\partial h}{\partial\theta}(x) = 0$ for all $x \in \mathbf{R}^k$. Hence, we see that (1) holds if $\sigma^2 = 0$. So suppose that $\sigma^2 > 0$ and set $\phi_z(t) := h(z + t\theta)$ for all $t \in \mathbf{R}$ and all $z \in \mathbf{R}^k$. Then ϕ_z is differentiable on $\mathbf{R}$ with derivative $\phi_z'(t) = \frac{\partial h}{\partial\theta}(z + t\theta)$. Let us define

$$B = \{z \in \mathbf{R}^k \mid \textstyle\int_{\mathbf{R}} |\frac{\partial h}{\partial\theta}(z+t\theta)| e^{-(\sigma t)^2/2} dt < \infty\}$$

and set $V_0 := \sigma^{-2}\,U_0$ and $V := (V_1, \ldots, V_k)$ where $V_i := U_i - \theta_i\,V_0$ for $i = 1, \ldots, k$. Then we have $V_0 \sim N(0, \sigma^{-2})$ and $E(V_0 V_i) = 0$ for all $1 \le i \le k$. Since $(V_0, V_1, \ldots, V_k)$ is Gaussian with mean zero, we see that V_0 and V are independent and since $U = V + V_0\theta$ and $\frac{\partial h}{\partial\theta}(U) \in L^1(P)$, we see that $P(V \in B) = 1$.

Let $z \in B$ be given and set $\psi(t) := -e^{-(\sigma t)^2/2}$ for all $t \in \mathbf{R}$. Then we have $\psi'(t) = \sigma^2 t\, e^{-(\sigma t)^2/2}$ and $\psi_*(t) = e^{-(\sigma t)^2/2}$ where $\psi_*(t)$ is defined as in Lemma 2.4. Since $z \in B$, we see that ϕ_z' is locally λ-integrable on $\mathbf{R}$. So by Theorem IX.4.5 p. 271 and Theorem IX.7.4 p. 284 in [15] we have that ϕ_z is absolutely continuous on $\mathbf{R}$ with Lebesgue derivative ϕ_z' and that (ϕ_z, ψ) satisfies condition

(1) in Lemma 2.5. Hence, we have that $\phi'_z\,\psi$ and $\phi_z\,\psi'$ are λ_1-integrable and

$$\int_{\mathbf{R}} h(z+t\theta)\sigma^2 t e^{-(\sigma t)^2/2} dt = \int_{\mathbf{R}} \phi_z(t)\psi'(t)\,dt = -\int_{\mathbf{R}} \phi'_z(t)\psi(t)\,dt$$
$$= \int_{\mathbf{R}} \tfrac{\partial h}{\partial \theta}(z+t\theta)e^{-(\sigma t)^2/2} dt.$$

Recall that V_0 and V are independent with $U = V + V_0\theta$ and $V_0 \sim N(0,\sigma^{-2})$. Since $U_0 h(U) \in L^1(P)$ and $\frac{\partial h}{\partial \theta}(U) \in L^1(P)$, we have $P(V \in B) = P_V(B) = 1$ and

$$\begin{aligned} E(U_0 h(U)) &= \sigma^2 E\{V_0 h(V + V_0\theta)\} \\ &= \tfrac{\sigma}{\sqrt{2\pi}} \int_B P_V\,(dz) \int_{\mathbf{R}} h(z+t\theta)\sigma^2 t e^{-(\sigma t)^2/2} dt \\ &= \tfrac{\sigma}{\sqrt{2\pi}} \int_B P_V(dz) \int_{\mathbf{R}} \tfrac{\partial h}{\partial \theta}(z+t\theta)e^{-(\sigma t)^2/2}\, dt \\ &= E(\tfrac{\partial h}{\partial \theta}(V + V_0\theta)) = E(\tfrac{\partial h}{\partial \theta}(U)) \end{aligned}$$

which proves the lemma. □

Lemma 2.7. *Let $Z = (Z_1, \ldots, Z_n)$ be an n-dimensional Gaussian random vector with mean zero and a non-zero covariance matrix $\Sigma^Z = \{\sigma^Z_{ij}\}$. Let λ denote the largest eigenvalue of Σ^Z and let ν denote the multiplicity of the eigenvalue λ. Let r denote the rank of Σ^Z and let $\phi : [0,\infty) \to [0,\infty)$ be an essentially increasing Borel function (see the remark below). Then $\lambda > 0$ and $1 \le \nu \le r \le n$ and we have*

(1) $\displaystyle\int_0^\infty t^{r-1}\phi(t)e^{-\frac{1}{2\lambda}t^2}\,dt < \infty \;\Rightarrow\; E\phi(\|Z\|) < \infty.$

(2) $\displaystyle E\phi(\|Z\|) < \infty \;\Rightarrow\; \int_0^\infty (1+t)^{\nu-1}\phi(t)e^{-\frac{1}{2\lambda}t^2}\,dt < \infty.$

(3) $E\phi(\|Z\|) < \infty \;\Rightarrow\; \exists c > 0$ *so that* $\phi(t) \le c(1+t)^{2-\nu}\exp\bigl(\frac{1}{2\lambda}t^2\bigr)\ \forall t \ge 0.$

Remark. We say that $\phi : [0,\infty) \to [0,\infty)$ is *essentially increasing* if there exists $C \ge 0$ such that $\phi(s) \le C\,(1 + \phi(t))$ for all $0 \le s \le t$.

Proof. Since $\Sigma^Z \ne 0$, we have $\lambda > 0$ and $1 \le \nu \le r \le n$. Let $v_1, \ldots, v_n \in \mathbf{R}^n$ be an orthonormal basis of eigenvectors of Σ^Z ordered such that $\lambda_1 \ge \lambda_2 \ge \cdots \ge \lambda_n$ where λ_i is the eigenvalue associated to v_i. Then we have $\lambda = \lambda_i \ge \lambda_r > 0$ for $1 \le i \le \nu$ and $\lambda_i = 0$ for $r < i \le n$. Set $U_i := \lambda_i^{-1/2}\langle Z, v_i\rangle$ for $i = 1, \ldots, r$. Then $U_1, \ldots, U_r$ are independent $N(0,1)$-distributed random variables such that $Z = \sum_{i=1}^r \lambda_i^{1/2} U_i\, v_i$ a.s. and $||Z||^2 = \sum_{i=1}^r \lambda_i U_i^2$ a.s. Let $1 \le d \le r$ be a given integer and set $U^d = (U_1, \ldots, U_d)$.

Recall that $2\pi^{d/2}\,\Gamma(\frac{d}{2})^{-1}$ is the $(d-1)$-dimensional volume of the $(d-1)$-dimensional unit sphere $\{u \in \mathbf{R}^d \mid ||u|| = 1\}$; see (8.43.9) p. 60 in vol. 2 of [3]. So by (8.24.1) p. 27 in vol. 2 of [3] with $T(x) = ||x||$, we have

$$E\phi(\sqrt{\lambda}\|U^d\|) = (2\pi)^{-\frac{d}{2}} \int_{\mathbf{R}^d} \phi(\sqrt{\lambda}\|x\|) e^{-\frac{1}{2}\|x\|^2} dx$$
$$= \frac{2}{2^{d/2}\Gamma(\frac{d}{2})} \int_0^\infty t^{d-1} \phi(t\sqrt{\lambda}) e^{-\frac{1}{2}t^2}\, dt = \frac{2}{(2\lambda)^{d/2}\Gamma(\frac{d}{2})} \int_0^\infty s^{d-1}\phi(s) e^{-\frac{1}{2\lambda}s^2}\, ds.$$

Since $||Z||^2 = \sum_{i=1}^r \lambda_i U_i^2$, we have $\lambda\,||U^\nu||^2 \le ||Z||^2 \le \lambda\,||U^r||^2$ and since ϕ is essentially increasing, there exists a constant $C > 0$ such that $0 \le \phi(s) \le C\,(1+\phi(t))$ for all $0 \le s \le t$. In particular, we have

$$\phi(\sqrt{\lambda}\|U^\nu\|) \le C(1+\phi(\|Z\|))\ ,\ \phi(\|Z\|) \le C(1+\phi(\sqrt{\lambda}\|U^r\|))$$

and since $\phi(t) \le C(1+\phi(1))$ for $0 \le t \le 1$ and $(1+t)^{\nu-1} \le 2^{\nu-1}\, t^{\nu-1}$ for $t \ge 1$, we see that (1) and (2) follow from the equality above. Since $\phi(s) \le C\,(1+\phi(t))$ for all $t \ge s$, we have

$$\phi(s)(1+s)^{\nu-1} \int_s^\infty e^{-\frac{1}{2\lambda}t^2}\, dt \le C \int_0^\infty (1+t)^{\nu-1}(1+\phi(t)) e^{-\frac{1}{2\lambda}t^2} dt$$

and so we see that (3) follows from (2) and Exercise 2.51 p. 148 in vol. 1 of [3]. □

Theorem 2.8 (Slepian's inequality). *Let $k \ge 1$ be an integer and let $X = (X_1, \ldots, X_k)$ and $Y = (Y_1, \ldots, Y_k)$ be Gaussian random vectors with zero means and covariance matrices $\Sigma^X = \{\sigma_{ij}^X\}$ and $\Sigma^Y = \{\sigma_{ij}^Y\}$. Let $K \subseteq \mathbf{R}^k$ be a bounded, starshaped Borel set with non-empty interior. Let $\phi : [0,\infty) \to [0,\infty)$ be an essentially increasing Borel function and let $f : \mathbf{R}^k \to \mathbf{R}$ be a locally λ_k-integrable Borel function satisfying*

(1) $P(X \in C_{\text{ap}}^K(f)) = 1 = P(Y \in C_{\text{ap}}^K(f))$.

(2) $E\phi(\|X\|) + E\phi(\|Y\|) < \infty$ *and* $\sup_{y\in K} |f(x+\delta y)| \le c\,\phi(\|x\|)\ \forall x \in \mathbf{R}^k$

for some positive numbers $c, \delta > 0$. Then $E|f(X)| < \infty$ and $E|f(Y)| < \infty$ and we have

(3) $\sum_{i=1}^k \sum_{j=1}^k (\sigma_{ij}^X - \sigma_{ij}^Y)\, \partial_{ij} f \ge 0 \ \Rightarrow\ Ef(Y) \le Ef(X)$.

Set $e = (1, 1, \ldots, 1)$ and suppose that there exist functions $h : \mathbf{R}^k \to \mathbf{R}$ and $\psi : \mathbf{R} \to \mathbf{R}$ such that $f(x+te) = h(x) + \psi(t)$ for all $(t,x) \in \mathbf{R} \times \mathbf{R}^k$. Set $\pi_{ij}^X = E(X_i - X_j)^2$ and $\pi_{ij}^Y = E(Y_i - Y_j)^2$ for all $i, j \in [k]$. Then we have

(4) $\sum_{i=1}^k \sum_{j=1}^k (\pi_{ij}^Y - \pi_{ij}^X)\partial_{ij} f = 2 \sum_{i=1}^k \sum_{j=1}^k (\sigma_{ij}^X - \sigma_{ij}^Y)\partial_{ij} f$.

(5) $\sum_{i=1}^k \sum_{j=1}^k (\pi_{ij}^Y - \pi_{ij}^X)\partial_{ij} f \ge 0 \ \Rightarrow\ Ef(Y) \le Ef(X)$.

Proof. Since K is starshaped, we have $0 \in K$. Hence, by (2), we have $|f(x)| \le c\,\phi(||x||)$ and $E|f(X)| < \infty$ and $E|f(Y)| < \infty$.

So suppose that $\sum_{i=1}^k \sum_{j=1}^k \theta_{ij}\, \partial_{ij} f \ge 0$, where $\theta_{ij} := \sigma_{ij}^X - \sigma_{ij}^Y$ and let me show that $Ef(Y) \le Ef(X)$. Without loss of generality we may assume that

X and Y are independent. If $X = Y = 0$ a.s. then (3) holds trivially. So let us assume that $P(X \neq 0) + P(Y \neq 0) > 0$ and set $\tilde{\phi}(t) := \sup_{s\in[0,t]} \phi(s)$. Then $\tilde{\phi}$ is increasing with $\phi \leq \tilde{\phi}$ and since ϕ is essentially increasing, there exists $C > 0$ such that $\tilde{\phi}(t) \leq C\,(1 + \phi(t))$ for all $t \geq 0$. Let $c, \delta > 0$ be chosen according to (2). Replacing (ϕ, K) by $(c\,\tilde{\phi}, \delta K)$ we may, without loss of generality, assume that ϕ is increasing and that $|f(x+y)| \leq \phi(||x||)$ for all $x \in \mathbf{R}^k$ and all $y \in K$. Let λ_X and λ_Y denote the largest eigenvalues of Σ_X and Σ_Y, respectively, and set $\lambda = \max(\lambda_X, \lambda_Y)$ and $\kappa(x) = \phi(||x||)$ for $x \in \mathbf{R}^k$. Then $\lambda > 0$ and by (2) and Lemma 2.7 there exists $b > 0$ such that $\phi(t) \leq b(1+t)\, \exp(\frac{1}{2\lambda}\, t^2)$ for all $t \geq 0$.

Let $0 < r < 1$ and $n \in \mathbf{N}$ be fixed for a while. Since K has non-empty interior, there exists $g \in C^\infty_{\circ\circ}(\mathbf{R}^k)$ such that $g \geq 0$, $\{g \neq 0\} \subseteq K$ and $\int_K g(y)\, dy = 1$. Set $f_n(x) = \int_{\mathbf{R}^k} f(x + \frac{y}{n})\, g(y)\, dy$ for $x \in \mathbf{R}^k$ (see Lemma 2.2). By Lemma 2.2 we have that $f_n \in C^\infty_\kappa(\mathbf{R}^k)$ and that (f_n) satisfies (1)–(6) in Lemma 2.2. Since $0 < r < 1$ and $\phi(rt) \leq b(1+rt)\, \exp(\frac{r^2}{2\lambda} t^2)$ we have

$$\text{(i)}\quad A := E\left\{\left(1 + \sqrt{\|X\|^2 + \|Y\|^2}\right)\phi\left(r\,\sqrt{\|X\|^2 + \|Y\|^2}\right)\right\} < \infty.$$

Let $x, y \in \mathbf{R}^k$ be given and set $U_{x,y}(t) = r(t^{1/2}\, x + (1-t)^{1/2}\, y)$ and $V_{x,y}(t) = f_n(U_{x,y}(t))$ for $t \in [0,1]$. Then $U_{x,y}$ and $V_{x,y}$ are continuous on $[0,1]$ and continuously differentiable on $(0,1)$ with derivatives

$$\begin{aligned} U'_{x,y}(t) &= \tfrac{r}{2}(t^{-1/2}x - (1-t)^{-1/2}y) = \tfrac{r}{2\sqrt{t(1-t)}}((1-t)^{1/2}x - t^{1/2}y) \\ V'_{x,y}(t) &= \langle U'_{x,y}(t), \nabla f_n(U_{x,y}(t))\rangle \end{aligned}$$

for all $0 < t < 1$. By the Cauchy-Schwartz inequality, we have

$$\begin{aligned} \|U_{x,y}(t)\| &\leq rt^{1/2}\|x\| + r(1-t)^{1/2}\|y\| \leq r\sqrt{\|x\|^2 + \|y\|^2} \quad \forall 0 \leq t \leq 1 \\ \|U'_{x,y}(t)\| &\leq \tfrac{r}{2\sqrt{t(1-t)}}\sqrt{\|x\|^2 + \|y\|^2} \quad \forall 0 < t < 1. \end{aligned}$$

In particular, we see that $V_{x,y}$ is absolutely continuous on $[0,1]$ and so we have

$$\text{(ii)}\quad f_n(rx) - f_n(ry) = V_{x,y}(1) - V_{x,y}(0) = \int_0^1 V'_{x,y}(t)\, dt.$$

Let $0 < t < 1$ be given and let us define $U(t) = (U_1(t), \ldots, U_k(t)) := U_{X,Y}(t)$ and $V'(t) := V'_{X,Y}(t)$. Since $V'(t) = \langle U'(t), \nabla f_n(U(t))\rangle$, we have

$$\begin{aligned} &\|U(t)\| \leq r\sqrt{\|X\|^2 + \|Y\|^2}\;,\;\; |V'(t)| \leq \|U'(t)\| \cdot \|\nabla f_n(U(t))\| \\ &\|U'(t)\| \leq \tfrac{r}{2\sqrt{t(1-t)}}\sqrt{\|X\|^2 + \|Y\|^2}. \end{aligned}$$

Set $\vartheta_j = \frac{r^2}{2}\,(\theta_{1j}, \ldots, \theta_{kj})$ and $h_j(x) = \frac{\partial f_n}{\partial x_j}(x)$ for $x \in \mathbf{R}^k$ and $1 \leq j \leq k$. Since $f_n \in C^\infty_\kappa(\mathbf{R}^k)$, we have $h_j \in C^\infty_\kappa(\mathbf{R}^k)$ and since $|h_j(x)| \leq ||\nabla f_n(x)||$ and $\kappa(x) =$

$\phi(\|x\|)$, we have

$$|U_j'(t)h_j(U(t))| \le \|U'(t)\| \cdot \|\nabla f_n(U(t))\| \le \|\nabla f_n\|_\kappa \cdot \|U'(t)\| \cdot \phi(\|U(t)\|)$$
$$|\tfrac{\partial h_j}{\partial \vartheta_j}(U(t))| \le \|\vartheta_j\| \cdot \|\nabla h_j(U(t))\| \le \|\vartheta_j\| \cdot \|\nabla h_j\|_\kappa \cdot \phi(\|U(t)\|).$$

Hence, by (i) we see that $U_j'(t)\, h_j(U(t))$, $\frac{\partial h_j}{\partial \vartheta_j}(U(t))$ and $V'(t)$ are P-integrable and

$$\int_0^1 E|V'(t)|\, dt \le A\|\nabla f_n\|_\kappa \int_0^1 \tfrac{1}{\sqrt{t(1-t)}} dt < \infty.$$

So by (ii) and the Fubini-Tonelli theorem, we have

(iii) $Ef_n(rX) - Ef_n(rY) = \displaystyle\int_0^1 EV'(t)\, dt.$

Since X and Y are independent Gaussian random vector with zero means, we see that $(U_j'(t), U_1(t), \dots, U_k(t))$ is a $(k+1)$-dimensional Gaussian random vector with zero mean and $E(U_j'(t)\, U_i(t)) = \frac{r^2}{2}\, \theta_{ij}$. So by Lemma 2.6 we have

$$E\Big\{U_j'(t) h_j(U(t))\Big\} = E\Big\{\tfrac{\partial h_j}{\partial \vartheta_j}(U(t))\Big\} = \tfrac{r^2}{2} E\bigg\{\sum_{i=1}^k \theta_{ij}\, \tfrac{\partial^2 f_n}{\partial x_i \partial x_j}(U(t))\bigg\}$$

and by Proposition 2.3, we have $\sum_{i=1}^k \sum_{j=1}^k \theta_{ij}\, \frac{\partial^2 f_n}{\partial x_i \partial x_j}(x) \ge 0$ for all $x \in \mathbf{R}^k$. Since

$$V'(t) = \langle U'(t), \nabla f_n(U(t)\rangle = \sum_{j=1}^k U_j'(t) h_j(U(t))$$

we see that $EV'(t) \ge 0$ for all $0 < t < 1$ and so by (iii) we have $Ef_n(rY) \le Ef_n(rX)$ for all $n \in \mathbf{N}$ and all $0 < r < 1$. Since ϕ is increasing and $\sup_{y \in K} |f(x+y)| \le \phi(||x||)$, we have $|f_n(rx)| \le \phi(||x||)$ for all $0 < r \le 1$, all $x \in \mathbf{R}^k$ and all $n \in \mathbf{N}$; see Lemma 2.2. So by (2), continuity of f_n and Lebesgue's convergence theorem, we have

$$Ef_n(Y) = \lim_{r \uparrow 1} Ef_n(rY) \le \lim_{r \uparrow 1} Ef_n(rX) = Ef_n(X)$$

for all $n \ge 1$. By (1) and Lemma 2.2, we have $f_n(x) \to f(x)$ $(P_X + P_Y)$-a.s. and recall that $|f_n(x)| \le \phi(\|x|)$. So by (2) and Lebesgue's convergence theorem, we have

$$Ef(Y) = \lim_{n \to \infty} Ef_n(Y) \le \lim_{n \to \infty} Ef_n(X) = Ef(X)$$

which completes the proof of (3).

Suppose that there exist functions $h : \mathbf{R}^k \to \mathbf{R}$ and $\psi : \mathbf{R} \to \mathbf{R}$ such that $f(x + te) = h(x) + \psi(t)$ for all $(t, x) \in \mathbf{R} \times \mathbf{R}^k$. Note that $\pi_{ii}^X = \pi_{jj}^Y = 0$ and

$$\pi_{ij}^Y - \pi_{ij}^X = 2(\sigma_{ij}^X - \sigma_{ij}^Y) + (\sigma_{ii}^Y - \sigma_{ii}^X) + (\sigma_{jj}^Y - \sigma_{jj}^X)$$

Hence, we see (4)–(5) follow from (3) and Lemma 2.4. □

Remark 2.9. (a): Condition (1) is a weak smoothness restriction on f. Note that condition (1) holds if f is right or left continuous and since $\mathbf{R}^k \setminus C_{\text{ap}}^K(f)$ is a λ_k-null set, we see that (1) holds if Σ^X and Σ^Y are non-singular. However, Example A in introduction shows that some smoothness condition on f is needed.

(b): Condition (2) is a growth condition on f. Let $a, \epsilon > 0$ be positive numbers and let $\psi : [0, \infty) \to [0, \infty)$ be an essentially increasing function satisfying $E\psi(\epsilon + ||X||) + E\psi(\epsilon + ||Y||) < \infty$ and $|f(x)| \leq a\,\psi(||x||)$ for all $x \in \mathbf{R}^k$. Since K is bounded and ψ is essentially increasing, it follows easily that f satisfies condition (2) with $\phi(t) := \psi(\epsilon + t)$.

(c): Let λ_X denote the largest eigenvalues of Σ^X, let ν_X denote the multiplicity of λ_X and let r_X denote the rank of the covariance matrix Σ^X. Let $p > 0$ be a given number and set $\phi(t) = (1+t)^{-p}\, e^{\frac{1}{2\lambda}\, t^2}$ for all $t \geq 0$. By Lemma 2.7, we have $p > r_X \ \Rightarrow E\phi(||X||) < \infty \ \Rightarrow \ p > \nu_X$. Since, the Slepian inequality implicitly requires finiteness of $E|f(X)|$ and $E|f(Y)|$, we see that the growth condition (2) is close to be optimal.

(d): Let $\varphi : [0, \infty) \to \mathbf{R}$ be an increasing convex function and set $Q(x) := \max_{1 \leq i,j \leq k} |x_i - x_j|$ and $f(x) = \varphi(Q(x))$ for all $x = (x_1, \ldots, x_k) \in \mathbf{R}^k$. Fernique (see Theorem 2.1.2 p. 18 in [2]) has shown that $\pi_{ij}^Y \leq \pi_{ij}^X \ \forall\, i, j \in [k]$ implies $Ef(Y) \leq Ef(X)$. Note that $\pi_{ii}^X = \pi_{ii}^Y = 0$ and $f(x + te) = f(x)$ for all $(t, x) \in \mathbf{R} \times \mathbf{R}^k$ and in Corollary 4.5 below, we shall see that $\partial_{ij} f \leq 0$ for all $i \neq j$. Hence, we see that (5) is an extension of Fernique's version of Slepian's inequality.

3. Integral orderings

In this section we shall study an extension of the integral ordering to the set of finitely additive contents. Let $(S, \mathcal{A}, \mu)$ be a *content space*; i.e., $\mathcal{A} \subseteq 2^S$ is an algebra on the set S and $\mu : \mathcal{A} \to [0, \infty]$ is a finitely additive set function satisfying $\mu(\emptyset) = 0$. If $D \subseteq S$, we let $\mu^*(D) = \inf_{A \in \mathcal{A}, A \supseteq D} \mu(A)$ denote *the outer content* of D. We let $TM(\mu)$ denote the set of all *totally μ-measurable* real-valued functions (see Def. III.2.10 p. 106 in [1]), and we let $L^1(\mu)$ denote the set of all μ-integrable functions (see Def. III.2.17 p. 112 in [1]). If $f, f_1, f_2, \ldots \in \mathbf{R}^S$, we write $f_n \to^\mu f$ if $\mu^*(|f - f_n| > \varepsilon) \to 0$ for all $\varepsilon > 0$ (see Lem. III.2.7 p. 104 in [1]). If $f, g : S \to \bar{\mathbf{R}}$, we write $f \leq g$ μ-a.e. if $\mu^*(f > \epsilon + g) = 0$ for all $\epsilon > 0$. Note that $\mu^*(f > g) = 0$ implies $f \leq g$ μ-a.e. and the converse implication holds if μ is a measure. If $f : S \to \bar{\mathbf{R}}$, we let

$$\int^* f d\mu := \inf\{\textstyle\int_S \phi\, d\mu \mid \phi \in L^1(\mu), f \leq \phi \ \mu\text{ -a.e. }\} \quad (\inf \emptyset := \infty)$$
$$\int_* f\, d\mu := \sup\{\textstyle\int_S \phi d\mu \mid \phi \in L^1(\mu), \phi \leq f \ \mu\text{ -a.e. }\} \quad (\sup \emptyset := -\infty)$$

denote *the upper and lower μ-integral* of f; see [4]. We say that $\Phi \subseteq \mathbf{R}^S$ is *uniformly μ-integrable* if for every $\varepsilon > 0$ there exists $h \in L^1(\mu)$ such that $\int^* (|\phi| - h)^+ \, d\mu < \varepsilon$ for all $\phi \in \Phi$. If μ is a measure and $\Phi \subseteq L^1(\mu)$, then the reader easily verifies that uniform integrability as defined here coincides with the usual definition of uniform integrability; see for instance, (3.22.34) p. 187 in vol. 1 of [3]. Let $f, f_1, f_2, \ldots \in \bar{\mathbf{R}}^S$ be given functions. Then we write $f \leq_\mu \liminf f_n$ if and only if

$$\mu^*(A \cap \{f > t\}) \leq \liminf_{n\to\infty} \mu^*(A \cap \{f_n > s\}) \ \ \forall s < t \ \forall A \in \mathcal{A}.$$

We define $\mathcal{A}^\circ := \{A \in \mathcal{A} \mid \mu(A) < \infty\}$ and we write $\limsup f_n \leq_\mu f$ if and only if

$$\limsup_{n\to\infty} \mu^*(A \cap \{f_n > t\}) \leq \mu^*(A \cap \{f > s\}) \ \ \forall s < t \ \forall A \in \mathcal{A}^\circ.$$

Recall that $L^1(\mu) \subseteq TM(\mu)$ and if $\phi \in TM(\mu)$, then we have

(3.1) $\quad (f - f_n)^+ \to^\mu 0 \Rightarrow f \leq_\mu \liminf f_n \Rightarrow (f + \phi) \leq_\mu \liminf(f_n + \phi)$.

(3.2) $\quad (f_n - f)^+ \to^\mu 0 \Rightarrow \limsup f_n \leq_\mu f \Rightarrow \limsup(f_n + \phi) \leq_\mu (f + \phi)$.

Set $f_*(s) = \liminf_{n\to\infty} f_n(s)$ and $f^*(s) = \limsup_{n\to\infty} f_n(s)$ for all $s \in S$. If $(S, \mathcal{A}, \mu)$ is a measure space, then we have

(3.3) $\quad f_* \leq_\mu \liminf f_n$.

(3.4) $\quad$ If $f_1, f_2, \ldots$ are μ-measurable, then $\limsup f_n \leq_\mu f^*$.

If $\mathcal{L} \subseteq 2^S$, we let $W(S, \mathcal{L})$ denote the set of all functions $f : S \to \mathbf{R}$ such that for all $y > x$ there exists a set $L \in \mathcal{L} \cup \{\emptyset, S\}$ satisfying $\{f > y\} \subseteq L \subseteq \{f > x\}$, and we let $W_\circ^+(S, \mathcal{L})$ denote the set of all functions $f : S \to [0, \infty)$ such that for all $y > x > 0$ there exists a set $L \in \mathcal{L} \cup \{\emptyset\}$ satisfying $\{f > y\} \subseteq L \subseteq \{f > x\}$. We say that $\Phi \subseteq \mathbf{R}^S$ is $(\uparrow)$-*stable* if $\sup_{n\geq 1} \phi_n \in \Phi$ for every increasing sequence $(\phi_n) \subseteq \Phi$ satisfying $\sup_{n\geq 1} \phi_n(s) < \infty$ for all $s \in S$, and we say that Φ is *sequentially closed* if for every pointwise convergent sequence $\phi_1, \phi_2, \ldots \in \Phi$ we have $\phi \in \Phi$ where $\phi(s) = \lim_{n\to\infty} \phi_n(s)$ for all $s \in S$. We let $\mathcal{L}_\uparrow$ denote the set of all sets of the form $\cup_{n=1}^\infty L_n$ for some increasing sequence $(L_n) \subseteq \mathcal{L} \cup \{\emptyset\}$ and we let $\mathcal{L}_\downarrow$ denote the set of all sets of the form $\cap_{n=1}^\infty L_n$ for some decreasing sequence $(L_n) \subseteq \mathcal{L} \cup \{\emptyset\}$. If $f \in W(S, \mathcal{L})$, then we have

(3.5) $\quad \{f > t\} \in \mathcal{L}_\uparrow \ \ \forall t \geq \inf_{s\in S} f(s)\ ,\ \{f \geq t\} \in \mathcal{L}_\downarrow \ \forall t > \inf_{s\in S} f(s)$.

If $\mathcal{L}$ is a σ-algebra on S, then $W(S, \mathcal{L}) = M(S, \mathcal{L})$. If S is a topological space and $\mathcal{L}$ is the set of all open (closed) subsets of S, then $W(S, \mathcal{L})$ is the set of all lower (upper) semicontinuous functions. If $(S, \leq)$ is a proset and $\mathcal{L}$ is the set of all upper intervals, then $W(S, \mathcal{L}) = \mathrm{In}(S, \leq)$.

Lemma 3.1. *Let S be a non-empty set and let $\Phi \subseteq \mathbf{R}_+^S$ be a $(\uparrow)$-stable, convex cone. Let $J \subseteq \mathbf{R}$ be an interval with interior J° and let $h : J \to \mathbf{R}$ be a continuous, increasing, convex function such that $\inf_{x\in J} h(x) = 0$. Let $f : S \to J$ be a given function satisfying $(f - c1_S)^+ \in \Phi$ for all $c \in J^\circ$. Then we have $h \circ f \in \Phi$.*

Proof. Let $a = \inf J$ and $b = \sup J$ denote the endpoints of J. If $a = b$, then the lemma holds trivially. So suppose that $a < b$ and set $\theta_c(t) = (t - c)^+$ for all $t \in J$ and all $c \in \mathbf{R}$. Let Γ denote the convex cone generated by $\{\theta_c \mid c \geq a\}$ and let us define

$$h_0(t) = \sup\{\gamma(t) \mid \gamma \in \Gamma, \gamma(s) \leq h(s) \ \ \forall s \in J\} \ \ \forall t \in J.$$

Then h_0 is an increasing, convex function satisfying $0 \leq h_0(t) \leq h(t)$ for all $t \in J$ and since h_0 is lower semicontinuous on J, we have that h_0 is continuous on J. Let $a < c < b$ be given. Since h is convex, we have that the right-hand derivative $r := \lim_{x \downarrow c} \frac{h(x)-h(c)}{x-c}$ exists and is finite and satisfies $h(c) + r\,(x-c) \leq h(x)$ for all $x \in J$. Set $\gamma(x) := (h(c) + r\,(x-c))^+$ for all $x \in J$. If $r = 0$, we have $h(x) \geq h(c)$ for all $x \in J$ and since $\inf_{x \in J} h(x) = 0$, we have $h(c) = 0 = \gamma(x)$ for all $x \in J$. If $r > 0$, we have $\gamma = r\,\theta_u$ where $u = c - \frac{h(c)}{r}$ and since $\inf_{x \in J} h(x) = 0$ and $c + \frac{h(x)-h(c)}{r} \geq x \geq a$ for all $x \in J$, we have $u \geq a$. Hence, in either case, we have $\gamma \in \Gamma$ and $\gamma(x) \leq h(x)$ for all $x \in J$. Since $\gamma(c) = h(c) \geq h_0(c)$, we have $h_0(c) = h(c)$ for all $c \in J^\circ$ and so by continuity of h and h_0, we have $h = h_0$.

By Lindelöf's theorem there exist $\gamma_1, \gamma_2, \ldots \in \Gamma$ such that $h(x) = \sup_{n \geq 1} \gamma_n(x)$ for all $x \in J$. Set $h_n(x) = \max(\gamma_1(x), \ldots, \gamma_n(x))$ for $x \in J$ and $n \geq 1$. Note that Γ is the set of all increasing, continuous, convex, piecewise linear functions $\gamma : J \to \mathbf{R}$ satisfying $\inf_{x \in J} \gamma(x) = 0$. In particular, we see that $h_n \in \Gamma$ and that $h_n(x) \uparrow h(x)$ for all $x \in J$. By assumption, we have $\theta_c \circ f \in \Phi$ for all $c \in J^\circ$ and since $\theta_c = 0$ for $c \geq b$ and Φ is an $(\uparrow)$-stable convex cone, we have $\gamma \circ f \in \Phi$ for all $\gamma \in \Gamma$. In particular, we have $h_n \circ f \in \Phi$ for all $n \geq 1$ and since $h_n(f(s)) \uparrow h(f(s))$ for all $s \in S$, we have $h \circ f \in \Phi$ $\square$

Lemma 3.2. *Let $(S, \mathcal{A}, \mu)$ be a content space and let $f, f_1, f_2, \ldots : S \to \bar{\mathbf{R}}$ be given functions. If $\{f_n^+) \mid n \geq 1\}$ is uniformly μ-integrable, then we have*

(1) $\limsup f_n \leq_\mu f \Rightarrow \limsup\limits_{n\to\infty} \int^* f_n d\mu \leq \int^* f d\mu.$

If $\{f_n^-) \mid n \geq 1\}$ is uniformly μ-integrable, then we have

(2) $f \leq_\mu \liminf f_n \Rightarrow \int^* f d\mu \leq \liminf\limits_{n\to\infty} \int^* f_n d\mu.$

Proof. Suppose that $\limsup f_n \leq_\mu f$ and that $\{f_n^+ \mid n \geq 1\}$ is uniformly μ-integrable. If $\int^* f\,d\mu = \infty$, then (1) holds trivially. So suppose that $\int^* f\,d\mu < \infty$ and let $\phi \in L^1(\mu)$ be given function satisfying $f \leq \phi$ μ-a.e. Set $g = f - \phi$ and $g_n = f_n - \phi$. Let $\varepsilon > 0$ be given. Since $g_n^+ \leq f_n^+ + \phi^-$, we see that (g_n^+) is uniformly μ-integrable. Hence, there exists $\psi \in L^1(\mu)$ such that $\psi \geq 0$ and $\int^* (g_n^+ - \psi)^+\,d\mu < \frac{\varepsilon}{2}$ for all $n \geq 1$. Since $\psi \in L^1_+(\mu)$, there exist positive numbers $\delta, c > 0$ such that $\int_S (\psi \wedge \delta)\,d\mu < \frac{\varepsilon}{2}$ and $\int_S (\psi - c)^+\,d\mu < \frac{\varepsilon}{2}$. Since $\mu^*(\psi > \delta) < \infty$, there exists $F \in \mathcal{A}$ such that $\{\psi > \delta\} \subseteq F$ and $\mu(F) < \infty$. Let $n \geq 1$ be given. Since $S \setminus F \subseteq \{\psi \leq \delta\}$ and $g_n^+ \leq \psi + (g_n^+ - \psi)^+$, we have

$$\int^* 1_{S \setminus F}\, g_n^+ d\mu \leq \int^* (\psi \wedge \delta) d\mu + \int^* (g_n^+ - \psi) d\mu \leq \tfrac{\varepsilon}{2} + \tfrac{\varepsilon}{2} = \varepsilon.$$

Set $h_n = 1_F\, g_n^+$ and $Q_n(t) = \mu^*(h_n > t)$ for $t \in \mathbf{R}_+$ and $n \geq 1$. Let $t > 0$ be given. Since $f \leq \phi$ μ-a.e., we have $\mu^*(g > t) = 0$ and by (3.2) we have $\limsup g_n \leq_\mu g$. Since $\{h_n > t\} = F \cap \{g_n > t\}$ and $\mu(F) < \infty$, we have $Q_n(t) \to 0$ for all $t > 0$.

Let $n \geq 1$ be given. Then we have $0 \leq h_n \leq g_n^+ \leq (g_n^+ - \psi)^+ + \psi$ and so by Theorem 2.1.(7) in [4] we have

$$\int_0^\infty Q_n(t)\,dt = \int^* h_n d\mu \leq \int^* (g_n^+ - \psi)^+\,d\mu + \int_S \psi\,d\mu \leq \tfrac{\varepsilon}{2} + \int_S \psi\,d\mu.$$

Since $0 \le h_n \le g_n^+$, we have $(h_n - c)^+ \le (g_n^+ - \psi)^+ + (\psi - c)^+$ and so by Theorem 2.1.(7) in [4] we have

$$\int_c^\infty Q_n(t)dt = \int_0^\infty Q_n(t+c)dt = \int^* (h_n - c)^+ \, d\mu$$
$$\le \int^* (g_n^+ - \psi)^+ \, d\mu + \int_S (\psi - c)^+ d\mu \le \tfrac{\varepsilon}{2} + \tfrac{\varepsilon}{2} = \varepsilon.$$

Since $0 \le h_n \le g_n^+ \le (g_n^+ - \psi)^+ + \psi$, we have $h_n \wedge \delta \le (g_n^+ - \psi)^+ + (\psi \wedge \delta)$ and so by Theorem 2.1.(7) in [4] we have

$$\int_0^\delta Q_n(t)\, dt = \int^* (h_n \wedge \delta)\, d\mu \le \int^* (g_n^+ - \psi)^+ \, d\mu + \int^* (\psi^+ \wedge \delta)\, d\mu \le \varepsilon.$$

Hence, for every $\varepsilon > 0$ there exist positive numbers $c, \delta > 0$ satisfying

$$\sup_{n\ge1} \int_0^\infty Q_n(t)dt < \infty \ , \ \sup_{n\ge1} \int_c^\infty Q_n(t)dt \le \varepsilon \ , \ \sup_{n\ge1} \int_0^\delta Q_n(t)\, dt \le \varepsilon.$$

Since Q_n is decreasing, it follows easily that $\{Q_n \mid n \ge 1\}$ is uniformly λ_1-integrable and recall that $Q_n(t) \to 0$ for all $t > 0$. Hence, by the Dunford-Pettis theorem (see (3.23) p. 189 in [3]) we have

$$0 = \lim_{n\to\infty} \int_0^\infty Q_n(t)\, dt = \lim_{n\to\infty} \int^* h_n \, d\mu$$

and since $g_n^+ = h_n + 1_{S\setminus F}\, g_n^+$, we have

$$\textstyle\int^* f_n \, d\mu \le \int_S \phi \, d\mu + \int^* g_n^+ \, d\mu \le \int_S \phi \, d\mu + \int^* h_n \, d\mu + \int^* 1_{S\setminus F} g_n^+ \, d\mu.$$

Since $\int^* 1_{S\setminus F}\, g_n^+ \, d\mu \le \varepsilon$ for all $n \ge 1$, we have

$$\limsup_{n\to\infty} \textstyle\int^* f_n \, d\mu \le \int_S \phi \, d\mu + \varepsilon \quad \forall \varepsilon > 0.$$

Letting $\varepsilon \downarrow 0$, we see that $\limsup \int^* f_n \, d\mu \le \int_S \phi \, d\mu$ for all $\phi \in L^1(\mu)$ with $f \le \phi$ μ-a.e. Taking infimum over ϕ, we obtain (1).

Suppose that $f \le_\mu \liminf f_n$ and that $\{f_n^- \mid n \ge 1\}$ is uniformly integrable. Let $\varepsilon > 0$ be given. Then there exists $\phi \in L^1(\mu)$ such that $\phi \ge 0$ and $\int^* (f_n^- - \phi)^+ \, d\mu \le \varepsilon$ for all $n \ge 1$. Set $g = f + \phi$ and $g_n = f_n + \phi$ and let us define $Q(t) = \mu^*(g > t)$ and $Q_n(t) = \mu^*(g_n > t)$ for $t \in \mathbf{R}$ and $n \ge 1$. By (3.1), we have $Q(t) \le \liminf Q_n(s)$ for all $0 < s < t$ and so we have $Q(t) \le \liminf Q_n(t)$ for all $t \in C(Q) \cap (0, \infty)$. Since Q is decreasing, we have $\mathbf{R} \setminus C(Q)$ is at most countable. Hence, by Theorem 2.1.(7) in [4] and Fatou's lemma we have

$$\int^* g^+ \, d\mu = \int_0^\infty Q(t)\, dt \le \liminf_{n\to\infty} \int_0^\infty Q_n(t)\, dt = \liminf_{n\to\infty} \int^* g_n^+ \, d\mu.$$

Since $g_n^- = (-f_n - \phi)^+ \le (f_n^- - \phi)^+$, we have (see Thm. 2.1.(5) in [4])

$$\textstyle\int^* g_n^+ \, d\mu = \int^* g_n \, d\mu + \int_* g_n^- \, d\mu \le \varepsilon + \int^* g_n \, d\mu$$

and since $\phi \in L^1(\mu)$ and $f = g + \phi$, we have

$$\int^* f\,d\mu \leq \int^* g^+\,d\mu + \int_S \phi\,d\mu \leq \liminf_{n\to\infty}(\int^* g_n^+\,d\mu + \int_S \phi\,d\mu)$$
$$\leq \varepsilon + \liminf_{n\to\infty}(\int^* g_n\,d\mu + \int_S \phi\,d\mu) = \varepsilon + \liminf_{n\to\infty} \int^* f_n\,d\mu.$$

Letting $\varepsilon \downarrow 0$, we obtain (2). □

Theorem 3.3. *Let $(S, \mathcal{A}_1, \mu)$ and $(S, \mathcal{A}_2, \nu)$ be content spaces and let us define*

$$\Lambda = \{f \in \mathbf{R}^S \mid \int^* f\,d\mu \leq \int^* f\,d\nu\}\ ,\ \mathcal{L} = \{L \subseteq S \mid \mu^*(L) \leq \nu^*(L)\}.$$

Let $f, g \in \mathbf{R}^S$ be given functions. Then we have

(1) $f \in \Lambda$ *and* $g \in \Lambda \cap L^1(\nu) \ \Rightarrow\ af + g \in \Lambda \ \ \forall a \in \mathbf{R}_+$.

(2) $f \vee (-n1_S) \in \Lambda \ \ \forall n \in \mathbf{N} \ \Rightarrow\ f \in \Lambda$.

(3) $\nu(S) < \infty$, $\nu(S) \leq \mu(S)$ *and* $(f + n1_S)^+ \in \Lambda \ \ \forall n \in \mathbf{N} \ \Rightarrow\ f \in \Lambda$.

(4) $W_\circ^+(S, \mathcal{L}) \subseteq \Lambda$ *and if* $\mu(S) = \nu(S) < \infty$, *then* $W(S, \mathcal{L}) \subseteq \Lambda$.

(5) *If $f_1, f_2, \ldots \in \Lambda$ are given functions such that $(f_n^- \mid n \geq 1)$ is uniformly μ-integrable, $(f_n^+ \mid n \geq 1)$ is uniformly ν-integrable, $f \leq_\mu \liminf f_n$ and $\limsup f_n \leq_\nu f$, then we have $f \in \Lambda$.*

(6) *If μ is a measure and $\int_* f\,d\mu > -\infty$, then $\{h \in \Lambda \mid h \geq f\}$ is $(\uparrow)$-stable.*

Let $J \subseteq \mathbf{R}$ be an interval with interior J° such that $f(S) \subseteq J$ and let $G : J \to \mathbf{R}$ be an increasing, continuous, convex function such that $\inf_{x\in J} G(x) = 0$. If $(S, \mathcal{A}_1, \mu)$ and $(S, \mathcal{A}_2, \nu)$ and $(T, \mathcal{B}, \eta)$ are measure spaces and $\mathcal{A} := \mathcal{A}_1 \cap \mathcal{A}_2$, then we have

(7) *If μ, ν and η are σ-finite and $h \in L^1(\mu \otimes \eta) \cap L^1(\nu \otimes \eta)$ is a given function such that $h(\cdot, t) \in \Lambda\ \forall t \in T$ and $h(s, \cdot) \in L^1(\eta)\ \forall s \in S$, then we have $h^\eta \in \Lambda \cap L^1(\mu) \cap L^1(\nu)$ where $h^\eta(s) := \int_T h(s,t)\,\eta(dt)$.*

(8) *If $f \in M(S, \mathcal{A}_2)$ and $(f - c1_S)^+ \in \Lambda \ \ \forall c \in J^\circ$, then $G \circ f \in \Lambda$.*

(9) *If $\Phi \subseteq M(S, \mathcal{A})$ is sequentially closed, $\{\phi^- \mid \phi \in \Phi\}$ is uniformly μ-integrable and $\{\phi^+ \mid \phi \in \Phi\}$ is uniformly ν-integrable, then $\Lambda \cap \Phi$ is sequentially closed.*

Proof. (1): Let $f \in \Lambda$ and $g \in \Lambda \cap L^1(\nu)$ be given functions and let $a \geq 0$ be a nonnegative number. Then we have (see [4]):

$$\int^* (af + g) d\mu \leq a \int^* f\,d\mu \dot{+} \int^* g\,d\mu \leq a \int^* f\,d\nu + \int_S g\,d\nu \leq \int^* (af + g)\,d\nu$$

which proves (1).

(2) is an immediate consequence of Theorem 2.1.(6) in [4] and since

$$f \vee (-n1_S) = (f + n1_S)^+ - n1_S,$$

we see that (3) follows from (1) and (2).

(4): Let $f \in W_\circ^+(S, \mathcal{L})$ be given. Let $y > x > 0$ be given. Since $\emptyset \in \mathcal{L}$, there exists $L \in \mathcal{L}$ such that $\{f > y\} \subseteq L \subseteq \{f > x\}$. Hence, we have $\mu^*(f > y) \leq \mu^*(L) \leq \nu^*(L) \leq \nu^*(f > x)$ for all $0 < x < y$ and so we have $\mu^*(f > x) \leq \nu^*(f > x)$ for all $x \in \mathbf{R}_+ \setminus D$ where D is the set of all discontinuity points of $y \curvearrowright \mu^*(f > y)$. Since $y \curvearrowright \mu^*(f > y)$ is decreasing, we have that D is at most countable and so by Theorem 2.1.(7) in [4] we have

$$\int^* f\,d\mu = \int_0^\infty \mu^*(f>t)\,dt \le \int_0^\infty \nu^*(f>t)\,dt = \int^* f\,d\nu$$

which proves the first inclusion in (4). So suppose that $\mu(S)=\nu(S)<\infty$ and let $f\in W(S,\mathcal{L})$ be given. Since $\{\emptyset,S\}\in\mathcal{L}$, we have $(f+n1_S)^+\in W_\circ^+(S,\mathcal{L})\subseteq\Lambda$ for all $n\ge 1$ and so by (3) we have $f\in\Lambda$ which completes the proof of (4).

(5): Suppose that the hypotheses of (5) hold. Then we have $\int^* f_n\,d\mu \le \int^* f_n\,d\nu$ and so by Lemma 3.2 we have

$$\textstyle\int^* f\,d\mu \le \liminf\limits_{n\to\infty} \int^* f_n d\mu \le \liminf\limits_{n\to\infty} \int^* f_n d\nu \le \limsup\limits_{n\to\infty} \int^* f_n d\nu \le \int^* f\,d\nu$$

which proves (5).

(6): Suppose that μ is a measure and $\int_* f\,d\mu > -\infty$. Let $(h_n)\subseteq\Lambda$ be an increasing sequence such that $h_1\ge f$ and $h_n\uparrow h\in\mathbf{R}^S$. Since $\int_* f\,d\mu>-\infty$, we have $\int^* f^-\,d\mu<\infty$ and since $0\le h_n^-\le h_1^-\le f^-$, we see that $\{h_n^-\mid n\ge 1\}$ is uniformly μ-integrable. Since $h_n\le h$ and $h_n\in\Lambda$, we have $\int^* h_n\,d\mu\le\int^* h_n\,d\nu\le\int^* h\,d\nu$ for all $n\ge 1$ and by (3.3), we have $h\le_\mu \liminf h_n$. So by Lemma 3.2, we see that $\int^* h\,d\mu\le\int^* h\,d\nu$ which proves (6).

(7): Suppose that the hypotheses of (7) hold. Since $h\in L^1(\mu\otimes\eta)\cap L^1(\nu\otimes\eta)$, there exists a η-null set $N\subseteq T$ such that $h(\,\cdot\,,t)\in L^1(\mu)\cap L^1(\nu)$ for all $t\in T\setminus N$. Since $h(\cdot,t)\in\Lambda$, we have $\int_S h(s,t)\,\mu(ds)\le\int_S h(s,t)\,\nu(ds)$ for all $t\in T\setminus N$ and by the Fubini-Tonelli theorem we have $h^\eta\in L^1(\mu)\cap L^1(\nu)$ and

$$\textstyle\int_S h^\eta\,d\mu = \int_{T\setminus N}\eta(dt)\int_S h(s,t)\,\mu(ds) \le \int_{T\setminus N}\eta(dt)\int_S h(s,t)\,\nu(ds) = \int_S h^\eta d\nu.$$

Hence, we see that $h^\eta\in\Lambda\cap L^1(\mu)\cap L^1(\nu)$.

(8): Suppose that $f\in M(S,\mathcal{A}_2)$ and $(f-c1_S)^+\in\Lambda$ for all $c\in J^\circ$. Set $\Phi=\Lambda\cap M_+(S,\mathcal{A}_2)$. By (1) and (6), we have that Φ is an $(\uparrow)$-stable cone. Let $\phi,\psi\in\Phi$ be given and let me show that $\phi+\psi\in\Lambda$. If $\int^*(\phi+\psi)\,d\nu=\infty$, this is evident. So suppose that $\int^*(\phi+\psi)\,d\nu<\infty$. Since $\psi\ge 0$, we have $\int^*\phi\,d\nu<\infty$ and since ϕ is nonnegative and ν-measurable, we have $\phi\in L^1(\nu)$, So by (1) we see that $\phi+\psi\in\Phi$. Hence, we see that Φ is an $(\uparrow)$-stable convex cone containing $(f-c1_S)^+\in\Lambda$ for all $c\in J^\circ$ and so by Lemma 3.1 we have $G\circ f\in\Phi\subseteq\Lambda$.

(9): Suppose that the hypotheses of (9) hold and let $h\in\mathbf{R}^S$ and $(h_n)\subseteq\Lambda\cap\Phi\cap M(S,\mathcal{A})$ be given functions satisfying $h_n(s)\to h(s)$ for all $s\in S$. Since Φ is sequentially closed, we have $h\in\Phi\cap M(S,\mathcal{A})$. By (3.3)–(3.4), we have $h\le_\mu \liminf h_n$ and $h\le_\nu \limsup h_n$ and since $\{h_n^-\mid n\ge 1\}$ is uniformly μ-integrable and $\{h_n^+\mid n\ge 1\}$ is uniformly ν-integrable, we have $h\in\Lambda$ by (5). □

Theorem 3.4. *Let $\Lambda\subseteq M(\mathbf{R}^k,\mathcal{B}^k)$ be a non-empty set, let $\kappa:\mathbf{R}^k\to[0,\infty]$ be a Borel function and let μ and ν be Borel measures on $\mathbf{R}^k$ satisfying*

(1) $\phi\star g\in\Lambda\ \ \forall\phi\in\Lambda\cap L^1_{\rm loc}(\lambda_k)\ \forall g\in C^\infty_{\circ\circ}(\mathbf{R}^k)_+$ *with* $\displaystyle\int_{\mathbf{R}^k} g(y)dy=1.$

(2) $\kappa \in L^1(\mu) \cap L^1(\nu)$ *and* $\int_{\mathbf{R}^k} f d\mu \leq \int_{\mathbf{R}^k} f\, d\nu \quad \forall f \in \Lambda \cap C_\kappa^\infty(\mathbf{R}^k)$.

Let $K \subseteq \mathbf{R}^k$ be a bounded, starshaped, Borel set with non-empty interior and let $f \in \Lambda$ be a locally λ_k-integrable Borel function satisfying

(3) $\exists c, \delta > 0$ *so that* $\sup_{y \in K} |f(x + \delta y)| \leq c\, \kappa(x) \quad \forall x \in \mathbf{R}^k$.

(4) $\mu(\mathbf{R}^k \setminus C_{\rm ap}^K(f)) = 0 = \nu(\mathbf{R}^k \setminus C_{\rm ap}^K(f))$.

Then $f \in L^1(\mu) \cap L^1(\nu)$ and we have $\int_{\mathbf{R}^k} f\, d\mu \leq \int_{\mathbf{R}^k} f\, d\nu$.

Proof. Let $f \in \Lambda \cap L^1_{\rm loc}(\lambda_k)$ be a given function satisfying (3)–(4) and let $c, \delta > 0$ be chosen according to (3). Since K has non-empty interior, there exists a function $g \in C_{\circ\circ}^\infty(\mathbf{R}^k)_+$ such that $\{g \neq 0\} \subseteq \delta K$ and $\int_K g\, d\lambda_k = 1$. Let $n \geq 1$ be given and set $f_n(x) = \int_{\mathbf{R}^k} f(x + \frac{y}{n})\, g(y)\, dy$ for $x \in \mathbf{R}^k$; see Lemma 2.2. By (1) and Lemma 2.2, we have $f_n \in \Lambda \cap C_\kappa^\infty(\mathbf{R}^k)$. So by (2), we have $\int f_n\, d\mu \leq \int f_n\, d\nu$ for all $n \geq 1$. By (3) and Lemma 2.2.(3), we have $|f_n(x)| \leq c\, \kappa(x)$ for all $x \in \mathbf{R}^k$ and all $n \geq 1$ and by (4) and Lemma 2.2.(4), we have $f_n \to f$ μ-a.e. and ν-a.e. By (2) we have $\kappa \in L^1(\mu) \cap L^1(\nu)$ and so by Lebesgue's convergence theorem we have $f \in L^1(\mu) \cap L^1(\nu)$ and

$$\int_{\mathbf{R}^k} f d\mu = \lim_{n\to\infty} \int_{\mathbf{R}^k} f_n d\mu \leq \lim_{n\to\infty} \int_{\mathbf{R}^k} f_n\, d\nu = \int_{\mathbf{R}^k} f\, d\nu$$

which proves the theorem. □

4. Modular orderings

Let μ and ν be Borel probability measures on $\mathbf{R}^k$ such that $\mu \leq_{\rm bsm} \nu$. In the modern literature it is frequently claimed that this implies $\mu \leq_{\rm sm} \nu$; see for instance [10]. Theorem 4.8 below shows that we do have $\int f\, d\mu \leq \int f\, d\nu$ for a large class of unbounded, supermodular Borel functions, and that we do have $\mu \leq_{\rm sm} \nu$ if μ and ν are discrete measures with finitely many mass points. However, Example C of the introduction shows that this inequality may fail for some continuous, linear, modular function f satisfying $0 \leq f \leq 2$ μ-a.s. and ν-a.s. This shows that a closer glance at the supermodular ordering is needed. This section will be devoted to the study of supermodular functions and the modular orderings introduced in the introduction. Recall that $f : \mathbf{R}^k \to \mathbf{R}$ is supermodular if and only if $f(x \vee y) + f(x \wedge y) \leq f(x) + f(y)$ for all $x, y \in \mathbf{R}^k$. Here we shall use an equivalent definition: f is supermodular if and only if $\Delta_i^s \Delta_j^t f(x) \geq 0$ for all $1 \leq i \neq j \leq k$, all $x \in \mathbf{R}^k$ and all $s, t \in \mathbf{R}_+$; see [8], where Δ_i^s for $i \in [k]$ is the difference operator $\Delta_i^s f(x) = f(x + se_i) - f(x)$ and $\Delta_i^s \Delta_j^t$ is the second-order difference operator:

$$\Delta_i^s \Delta_j^t f(x) = f(x + se_i + te_j) - f(x + se_i) - f(x + te_j) + f(x).$$

Let $k \geq 1$ be a given integer and let $f : \mathbf{R}^k \to \mathbf{R}$ be a given function. If $i \in [k]$, we write $\Delta_i f \geq 0$ if and only if $\Delta_i^s f(x) \geq 0$ for all $s \in \mathbf{R}_+$ and all $x \in \mathbf{R}^k$.

If $i, j \in [k]$, we write $\Delta_{ij} f \geq 0$ if and only if $\Delta_i^s \Delta_j^t f(x) \geq 0$ for all $s, t \in \mathbf{R}_+$ and all $x \in \mathbf{R}^k$, and we write $\Delta_{ij} f \leq 0$ if and only if $\Delta_{ij}(-f) \geq 0$. If $x \in \mathbf{R}^k$ and $i \in [k]$, we let

$$f_i^x(t) = f(x + (t - x_i)e_i) = f(x_1, \dots, x_{i-1}, t, x_{i+1}, \dots, x_k) \quad \forall t \in \mathbf{R}$$

denote the partial function. Let "xxx" be a given property of a function of one variable (such as "increasing" or "continuous" or "differentiable"). If $i \in [k]$ and $f : \mathbf{R}^k \to \mathbf{R}$ is a function of k variables, we say that f has "xxx" *in the ith coordinate* if the partial functions f_i^x has "xxx" for all $x \in \mathbf{R}^k$. Note that f is increasing if and only if f is increasing in each coordinate and that we have

(4.1) $\Delta_j f \geq 0 \Leftrightarrow f$ is increasing in the jth coordinate.

(4.2) $\Delta_{ij} f \geq 0 \Leftrightarrow x \curvearrowright \Delta_i^s f(x)$ is increasing in the jth coordinate for all $s > 0$.

(4.3) f is supermodular if and only if $\Delta_{ij} f \geq 0$ for all $1 \leq i \neq j \leq k$; see [8].

(4.4) f is convex in the ith coordinate if and only if $\Delta_{ii} f \geq 0$ and f has the Baire property in the ith coordinate; see [20].

If $i \in [k]$ and $x \in \mathbf{R}^k$, we let

$$\overline{D}_i f(x) := \limsup_{u \to 0} \tfrac{f(x+ue_i) - f(x)}{u} \ , \quad \underline{D}_i f(x) := \liminf_{u \to 0} \tfrac{f(x+ue_i) - f(x)}{u}$$
$$\overline{D}_i^r f(x) := \limsup_{u \downarrow 0} \tfrac{f(x+ue_i) - f(x)}{u} \ , \quad \overline{D}_i^\ell f(x) := \limsup_{u \uparrow 0} \tfrac{f(x+ue_i) - f(x)}{u}$$
$$\underline{D}_i^r f(x) := \liminf_{u \downarrow 0} \tfrac{f(x+ue_i) - f(x)}{u} \ , \quad \underline{D}_i^\ell f(x) := \liminf_{u \uparrow 0} \tfrac{f(x+ue_i) - f(x)}{u}$$

denote *the right / left / upper / lower partial Dini derivatives* of f at x; see [15].

Proposition 4.1. *Let $f : \mathbf{R}^k \to \mathbf{R}$ and $\phi_1, \dots, \phi_k : \mathbf{R} \to \mathbf{R}$ be given functions and set $\phi(x) := (\phi_1(x_1), \dots, \phi_k(x_k))$ and $\zeta(x) := \prod_{i=1}^k \phi_i(x_i)$ for all $x = (x_1, \dots, x_k) \in \mathbf{R}^k$. Let $J \subseteq \mathbf{R}$ be an interval and let $\xi : J \to \mathbf{R}$ be an increasing and convex function. Let $i, j \in [k]$ be given integers. Then we have*

(1) *If $\phi_1, \dots, \phi_k$ are nonnegative, $i \neq j$ and ϕ_i and ϕ_j are both increasing (decreasing) on $\mathbf{R}$, then we have $\Delta_{ij} \zeta \geq 0$.*
(2) *If $\Delta_{ij} f \geq 0$, $i \neq j$ and ϕ_i and ϕ_j are both increasing (decreasing) on $\mathbf{R}$, then we have $\Delta_{ij}(f \circ \phi) \geq 0$.*
(3) *Let $h_1, \dots, h_n : \mathbf{R}^k \to \mathbf{R}$ and $g : \mathbf{R}^n \to \mathbf{R}$ be increasing functions and set $h(x) := (h_1(x), \dots, h_n(x))$ for all $x \in \mathbf{R}^k$. Then we have*

$$\Delta_{ij} h_\ell \geq 0 \text{ and } \Delta_{\ell m} g \geq 0 \quad \forall 1 \leq \ell, m \leq n \Rightarrow \Delta_{ij}(g \circ h) \geq 0.$$

(4) *If $f(\mathbf{R}^k) \subseteq J$ and $\Delta_{ij}(f \vee a) \geq 0 \ \forall a \in J$, then $\Delta_{ij}(\xi \circ f) \geq 0$.*
(5) *If $f(\mathbf{R}^k) \subseteq J$, f is increasing and $\Delta_{ij} f \geq 0$, then $\Delta_{ij}(\xi \circ f) \geq 0$.*

Proof. (1) and (2) are easy and well known. Let g and $h_1, \ldots, h_n$ be increasing functions such that $\Delta_{ij} h_\ell \geq 0$ and $\Delta_{\ell m} g \geq 0$ for all $1 \leq \ell, m \leq n$. Set $\psi(x) := g(h(x))$ for $x \in \mathbf{R}^k$. Let $u, v > 0$ and $x \in \mathbf{R}^k$ be given and let us define $y = x + ve_j$ and

$$\begin{aligned}
&x^0 = h(x)\ ,\ x^n = h(x + ue_i)\ ,\ y^0 = h(y)\ ,\ y^n = h(y + ue_i) \\
&x^\ell = (h_1(x + ue_i), \ldots, h_\ell(x + ue_i), h_{\ell+1}(x), \ldots, h_n(x)) \ \ \forall 1 \leq \ell < n \\
&y^\ell = (h_1(y + ue_i), \ldots, h_\ell(y + ue_i), h_{\ell+1}(y), \ldots, h_n(y)) \ \ \forall 1 \leq \ell < n.
\end{aligned}$$

Let $1 \leq \ell \leq n$ be given and set $u_\ell := \Delta_i^u h_\ell(x)$ and $v_\ell := \Delta_i^u h_\ell(y)$. Then we have $g(x^\ell) - g(x^{\ell-1}) = \Delta_\ell^{u_\ell} g(x^{\ell-1})$ and $g(y^\ell) - g(y^{\ell-1}) = \Delta_\ell^{v_\ell} g(y^{\ell-1})$. Since h_ℓ is increasing, we have $u^\ell \geq 0$ and recall that $\Delta_{ij} h_\ell \geq 0$. Hence, by (4.2) we have that $z \curvearrowright \Delta_i^u h_\ell(z)$ is increasing in the jth coordinate and so we have $0 \leq u_\ell \leq v_\ell$. Recall that $\Delta_{\ell m} g \geq 0$ for all $m = 1, \ldots, n$. Hence, by (4.2) we have that $z \curvearrowright \Delta_\ell^{u_\ell} g(z)$ is increasing on $\mathbf{R}^n$ and since $x \leq y$ and $h_1, \ldots, h_n$ are increasing, we have $x^{\ell-1} \leq y^{\ell-1}$. Thus, we have $\Delta_\ell^{u_\ell} g(x^{\ell-1}) \leq \Delta_\ell^{u_\ell} g(y^{\ell-1})$ and since $0 \leq u_\ell \leq v_\ell$ and g is increasing, we have $\Delta_\ell^{u_\ell} g(y^{\ell-1}) \leq \Delta_\ell^{v_\ell} g(y^{\ell-1})$. Hence, we have

$$\begin{aligned}
\Delta_i^u \psi(x) &= g(x^n) - g(x^0) = \textstyle\sum_{\ell=1}^n (g(x^\ell) - g(x^{\ell-1}) = \sum_{\ell=1}^n \Delta_\ell^{u_\ell} g(x^{\ell-1}) \\
&\leq \textstyle\sum_{\ell=1}^n \Delta_\ell^{v_\ell} g(y^{\ell-1}) = \sum_{\ell=1}^n (g(y^\ell) - g(y^{\ell-1}) = g(y^n) - g(y^0) = \Delta_i^u \psi(x + ve_j)
\end{aligned}$$

for all $x \in \mathbf{R}^k$ and all $u, v > 0$. In particular, we see that $x \curvearrowright \Delta_i^u \psi(x)$ is increasing in the jth coordinate for all $u > 0$ and so by (4.2) we conclude that $\Delta_{ij} \psi \geq 0$. Thus, (3) is proved.

(4): Suppose that $\Delta_{ij}(f \vee a) \geq 0$ for all $a \in J$. Let Φ_{ij} denote the set of all functions $F : \mathbf{R}^k \to \mathbf{R}$ such that $\Delta_{ij} F \geq 0$. Then Φ_{ij} is a pointwise closed, convex cone containing all constant functions and since $(f(x) - a)^+ = (f \vee a)(x) - a$, we see that $(f(\cdot) - a)^+ \in \Phi_{ij}$ for all $a \in J$. Since ξ is increasing and convex on J, there exist increasing, continuous, convex functions $\xi_1, \xi_2, \ldots : \mathbf{R} \to \mathbf{R}$ such that $\xi_m(t) \to \xi(t)$ for all $t \in J$ and $c_m := \inf_{t \in J} \xi_m(t) > -\infty$ for all $m \geq 1$. Then $\eta_m(t) = \xi_m(t) - c_m$ is an increasing, continuous, convex function on J with $\inf_{t \in J} \eta_m(t) = 0$. So by Lemma 3.1, we see that $\eta_m \circ f \in \Phi_{ij}$ for all $m \geq 1$ and since $\eta_m(t) + c_m = \xi_m(t) \to \xi(t)$, we have $\Delta_{ij}(\xi \circ f) \geq 0$.

(5): Suppose that f is increasing with $\Delta_{ij} f \geq 0$ and let $\xi_1, \xi_2, \ldots : \mathbf{R} \to \mathbf{R}$ be chosen as above. By (4.4), we have $\Delta_{11} \xi_m \geq 0$ and so by (3) applied with $n := 1$ and $(g, h_1) := (\xi_m, f)$, we see that $\Delta_{ij}(\xi_m \circ f) \geq 0$ for all $m \geq 1$ and since $\xi_m(f(x)) \to \xi(f(x))$, we have $\Delta_{ij}(\xi \circ f) \geq 0$. □

Proposition 4.2. *Let $f : \mathbf{R}^k \to \mathbf{R}$ be a supermodular function. If $i \in [k]$ and $s, t \in \mathbf{R}$, then we have*

(1) *$x \curvearrowright f_i^x(t) - f_i^x(s)$ is increasing on $\mathbf{R}^k$ if $s \leq t$ and decreasing on $\mathbf{R}^k$ if $t \leq s$.*

Let us define $\sigma(x,y) = \{i \in [k] \mid x_i < y_i\}$ *for* $x = (x_1, \dots, x_k) \in \mathbf{R}^k$ *and* $y = (y_1, \dots, y_k) \in \mathbf{R}^k$. *Let us define*

$$F_a(x) = f(x \vee a) - \sum_{i=1}^{n} f_i^a(x_i \vee a_i) \ , \ F^a(x) = f(x \wedge a) - \sum_{i=1}^{n} f_i^a(x_i \wedge a_i)$$

for all $a = (a_1, \dots, a_k) \in \mathbf{R}^k$ *and all* $x = (x_1, \dots, x_k) \in \mathbf{R}^k$. *If* $a \leq b$ *and* $x, y \in [a, b]$ *are given vectors, we have*

(2) $f(x) - f(y) \leq \sum_{i \in \sigma(x,y)} (f_i^a(x_i) - f_i^a(y_i)) + \sum_{i \in \sigma(y,x)} (f_i^b(x_i) - f_i^b(y_i))$.

(3) $|f(x) - f(y)| \leq \sum_{i=1}^{k} |f_i^a(x_i) - f_i^a(y_i)| + \sum_{i=1}^{k} |f_i^b(x_i) - f_i^b(y_i)|$.

(4) f *is modular if and only if there exist functions* $f_1, \dots, f_k : \mathbf{R} \to \mathbf{R}$ *such that* $f(u) = f_1(u_1) + \dots + f_k(u_k)$ *for all* $u = (u_1, \dots, u_k) \in \mathbf{R}^k$.

(5) F_a *is increasing and supermodular and* F^a *is decreasing and supermodular.*

Proof. (1): Let $1 \leq i \neq j \leq k$ and $s, t \in \mathbf{R}$ be given such that $s \leq t$. Then $u := t - s \geq 0$ and by (4.2)–(4.3) and supermodularity of f, we have that $x \curvearrowright \Delta_i^u f(x)$ is increasing in the jth coordinate. Since $f_i^x(t) - f_i^x(s) = \Delta_i^u f(x + (s - x_i)e_i)$, we see that $x \curvearrowright f_i^x(t) - f_i^x(s)$ is increasing in the jth coordinate for $j \neq i$ and since $x \curvearrowright f_i^x(t) - f_i^x(s)$ is constant in the ith coordinate, we see that $x \curvearrowright f_i^x(t) - f_i^x(s)$ is increasing if $s \leq t$. Interchanging s and t, we see that $x \curvearrowright f_i^x(t) - f_i^x(s)$ is decreasing if $t \leq s$.

(2)–(3): Let $x, y \in \mathbf{R}^k$ be given vectors and set $z_0 = x$, $z_n = y$ and $z_i = (y_1, \dots, y_i, x_{i+1}, \dots, x_k)$ for $1 \leq i < n$. Then we have

(i) $f(x) - f(y) = \sum_{i=1}^{n} (f(z_{i-1}) - f(z_i)) = \sum_{i=1}^{n} (f_i^{z_i}(x_i) - f_i^{z_i}(y_i))$.

Let $a, b \in \mathbf{R}^k$ be given vectors such that $x, y \in [a, b]$. Since $a \leq z_i \leq b$, we have $f_i^{z_i}(x_i) - f_i^{z_i}(y_i) \leq f_i^a(x_i) - f_i^a(y_i)$ for all $i \in \sigma(x, y)$ and $f_i^{z_i}(x_i) - f_i^{z_i}(y_i) \leq f_i^b(x_i) - f_i^b(y_i)$ for all $i \in \sigma(y, x)$. Since $f_i^{z_i}(x_i) - f_i^{z_i}(y_i) = 0$ for $i \notin \sigma(x, y) \cup \sigma(y, x)$, we see that (2) holds. (3) is an immediate consequence of (2).

(4): So suppose that f is modular. By (1), there exist functions $g_1, \dots, g_k : \mathbf{R} \to \mathbf{R}$ such that $f_i^x(t) - f_i^x(0) = g_i(t)$ for all $(t, x) \in \mathbf{R} \times \mathbf{R}^k$ and all $i \in [k]$. So by (i) with $y = (0, \dots, 0)$, we see that $f(x) = f(0, \dots, 0) + \sum_{i=1}^{k} g_i(x_i)$ which proves the "only if" in (4). The "if" part is evident.

(5): Let $x \leq y$ be given. Then we have $a \leq x \vee a \leq y \vee a$ and so by (2) with $b := y \vee a$, we have

$$f(x \vee a) - f(y \vee a) \leq \textstyle\sum_{i=1}^{k} (f_i^a(x_i \vee a_i) - f_i^a(y_i \vee a_i)).$$

Hence, we see that F_a is increasing. In the same manner we see that F^a is decreasing. By Proposition 4.1.(2) we see that $f(x \vee a)$ and $f(x \wedge a)$ are supermodular. So by (4) we see that F_a and F^a are supermodular. $\square$

Proposition 4.3. *Let* $f : \mathbf{R}^k \to \mathbf{R}$ *be a supermodular function. Let* $D \subseteq \mathbf{R}^k$ *be a given set satisfying* $\cup_{u \in D}[u, *] = \mathbf{R}^k = \cup_{u \in D}[*, u]$. *Let* $\mathcal{A}_1, \dots, \mathcal{A}_k$ *be* σ*-algebras* *on* $\mathbf{R}$*, let* μ_i *be a finite measure on* $(\mathbf{R}, \mathcal{A}_i)$ *for* $i \in [k]$ *and let* $\mathcal{A} = \mathcal{A}_1 \otimes \cdots \otimes \mathcal{A}_k$ *denote the product* σ*-algebra on* $\mathbf{R}^k$*. Let* $\tau_1, \ldots, \tau_k$ *be topologies on* $\mathbf{R}$ *and let* $\tau = \tau_1 \times \cdots \times \tau_k$ *denote the product topology on* $\mathbf{R}^k$*. If* $c = (c_i, \ldots, c_k)$ *is a given vector such that* c_i *admits a bounded* τ_i*-neighborhood for all* $i \in [k]$*, then we have*

(1) *If* f_i^u *is* $\mathcal{A}_i$*-measurable* $\forall\, u \in D\ \forall\, i \in [k]$*, then* f *is* $\mathcal{A}$*-measurable.*

(2) *If* f_i^u *is* τ_i*-continuous at* $c_i\ \ \forall\, u \in D\ \forall\, i \in [k]$*, then* f *is* τ*-continuous at* $c = (c_1, \ldots, c_k)$*.*

(3) *If* $f_i^u \in L^1(\mu_i)\ \forall\, u \in D\ \forall\, i \in [k]$*, then* $f_i^x \in L^1(\mu_i)\ \forall\, x \in \mathbf{R}^k\ \forall\, i \in [k]$*.*

Proof. Since $\mathbf{R}^k = \cup_{a \in D}\, [a, *]$ there exists $a^n = (a_1^n, \ldots, a_k^n) \in D$ for $n \geq 1$ such that $a_i^{n+1} < a_i^n \leq -n$ for all $n \in \mathbf{N}$ and all $i \in [k]$ and since $\mathbf{R}^k = \cup_{a \in D}\, [*, a]$ there exists $b^n = (b_1^n, \ldots, b_k^n) \in D$ for $n \geq 1$ such that $n \leq b_i^n < b_i^{n+1}$ for all $n \in \mathbf{N}$ and all $i \in [k]$. Set $C_n = [a^n, b^n]$ for all $n \geq 1$. Then we have $C_n \uparrow \mathbf{R}^k$ and by Proposition 4.2.(3) we have

(i) $|f(x) - f(y)| \leq \sum_{i=1}^{k} |f_i^{a^n}(x_i) - f_i^{a^n}(y_i)| + \sum_{i=1}^{k} |f_i^{b^n}(x_i) - f_i^{b^n}(y_i)|$

for all $x, y \in C_n$. Suppose that f_i^u is $\mathcal{A}_i$-measurable for all $u \in D$ and all $i \in [k]$. Since $f_i^{a^n}$ and $f_i^{b^n}$ are $\mathcal{A}_i$-measurable for all $n \in \mathbf{N}$ and all $i \in [k]$, it follows easily that f is $(\mathcal{A}_1 \otimes \cdots \otimes \mathcal{A}_k)$-measurable.

(2): Suppose that f_i^u is continuous at c_i for all $u \in D$ and all $i \in [k]$. By assumption, we have that c_i admits a bounded τ_i-neighborhood G_i for $i = 1, \ldots, k$. Since $a_i^n \leq -n < 0 < n \leq b_i^n$, there exists $q \in \mathbf{N}$ such that $G_i \subseteq [a_i^q, b_i^q]$ for all $i \in [k]$. Then $G := G_1 \times \cdots \times G_k$ is a τ-neighborhood of c such that $G \subseteq [a^q, b^q]$ and $f_i^{a^q}$ and $f_i^{b^q}$ are τ_i-continuous at c_i for all $i \in [k]$. So by (i) we see that f is τ-continuous at c.

(3): Suppose that $f_i^u \in L^1(\mu_i)$ for all $u \in D$ and all $i \in [k]$, Let $x \in \mathbf{R}^k$ and $i \in [k]$ be given. Then there exists $n \geq 1$ such that $x \in [a^n, b^n]$. So by Proposition 4.2 (1) we have

$$|f_i^x(t) - f_i^x(0)| \leq |f_i^{a^n}(t) - f_i^{a^n}(0)| + |f_i^{b^n}(t) - f_i^{b^n}(0)| \quad \forall t \in \mathbf{R}$$

and by (1), we see that f_i^x is μ_i-measurable. Since $(\mathbf{R}^k, \mathcal{A}_i, \mu_i)$ is a finite measure space and $f_i^{a^n}$ and f^{b^n} belong to $L^1(\mu_i)$, we see that $f_i^x \in L^1(\mu_i)$. $\square$

Theorem 4.4. *Let* $f : \mathbf{R}^k \to \mathbf{R}$ *be a given function and let* $i, j \in [k]$ *be given integers such that* f *is continuous in the* i*th coordinate. Then we have*

(1) *If* $\lambda_1(t \in \mathbf{R} \mid \overline{D}_i f(x + te_i) < 0) = 0$ *and* $\{t \in \mathbf{R} \mid \overline{D}_i f(x + te_i) = -\infty\}$ *is at most countable for all* $x \in \mathbf{R}^k$*, then* f *is increasing in the* i*th coordinate.*

Let $D_i^\diamond$ *denote one of the six Dini operators* $\overline{D}_i^r$, $\underline{D}_i^r$, $\overline{D}_i^\ell$, $\underline{D}_i^\ell$, $\overline{D}_i$ *or* $\underline{D}_i$*. and let us define*

$$I_{i,j}^x := \{t \in \mathbf{R} \mid D_i^\diamond f(x + te_j) > 0\}\ ,\ \ J_{i,j}^x := \{t \in \mathbf{R} \mid D_i^\diamond f(x + te_j) < 0\}$$

for all $x \in \mathbf{R}^k$. Suppose that $D_i^\diamond f(x)$ is finite for all $x \in \mathbf{R}^k$. If $J \subseteq \mathbf{R}$ is an interval containing $f(\mathbf{R}^k)$ and $\xi : J \to \mathbf{R}$ is an increasing convex function, then we have

(2) $\Delta_{ij} f \geq 0 \Leftrightarrow D_i^\diamond f(x)$ *is increasing in the jth coordinate.*

(3) *If $\Delta_{ij} f \geq 0$ and $t \curvearrowright f(x + te_j)$ is increasing on $I_{i,j}^x$ and decreasing on $J_{i,j}^x$ for all $x \in \mathbf{R}^k$, then $\Delta_{ij}(\xi \circ f) \geq 0$.*

(4) *If $\Delta_{ij} f \leq 0$ and $t \curvearrowright f(x + te_j)$ is decreasing on $I_{i,j}^x$ and increasing on $J_{i,j}^x$ for all $x \in \mathbf{R}^k$, then $\Delta_{ij}(\xi \circ f) \leq 0$.*

Proof. (1) follows from Theorem VI.7.3 p. 204 in [15]. So let $D_i^\diamond$ be one of the six Dini derivatives and suppose that $D_i^\diamond f(x)$ is finite for all $x \in \mathbf{R}^k$.

(2): Suppose that $\Delta_{ij} f \geq 0$ and let $x \in \mathbf{R}^k$ and $s > 0$ be given. By (4.2) we have that $t \curvearrowright \Delta_i^s f(x + te_j)$ is increasing on $\mathbf{R}$ and so we see that $D_i^\diamond f(x)$ is increasing in the jth coordinate. Conversely, suppose that $D_i^\diamond f(x)$ is increasing in the jth coordinate. Let $u > 0$ and $x \in \mathbf{R}^k$ be given and set $g(t) := \Delta_j^u f(x + te_i) = f(x + ue_j + te_i) - f(x + te_i)$ for all $t \in \mathbf{R}$. Since f is continuous in the ith coordinate, we have that g is continuous on $\mathbf{R}$ and since $D_i^\diamond f(x)$ is finite and increasing in the jth coordinate, it follows easily that we have

$$\overline{D}g(t) \geq D_i^\diamond f(x + ue_j + te_i) - D_i^\diamond f(x + te_i) \geq 0 \quad \forall t \in \mathbf{R}.$$

Hence, by Theorem VI.7.3 p. 204 in [15] we see that g is increasing; that is, $\Delta_i^u f$ is increasing in the jth coordinate and so by (4.2) we have $\Delta_{ij} f \geq 0$.

Let $\phi : \mathbf{R} \to \mathbf{R}$ be a continuously differentiable, increasing, convex function and set $h = \phi \circ f$. Let $x \in \mathbf{R}^k$ be given. Then we have

$$\Delta_i^s h(x) = \Delta_i^s f(x) \cdot \int_0^1 \phi'(f(x) + t \cdot \Delta_i^s f(x))\, dt \quad \forall s \in \mathbf{R}$$

and since f is continuous in the ith coordinate and ϕ' is increasing, nonnegative and continuous we have

(i) $D_i^\diamond h(x + te_j) = D_i^\diamond f(x + te_j)\phi'(f(x + te_j)) \quad \forall t \in \mathbf{R}$.

Suppose that $\Delta_{ij} f \geq 0$ and that $t \curvearrowright f(x + te_j)$ is increasing on $I_{i,j}^x$ and decreasing on $J_{i,j}^x$ for all $x \in \mathbf{R}^k$. By (2), we have that $t \curvearrowright D_i^\diamond f(x + te_j)$ is increasing on $\mathbf{R}$. Let $x \in \mathbf{R}^k$ be a given vector, let $s < t$ be given numbers and let me show that $D_i^\diamond h(x + se_j) \leq D_i^\diamond h(x + te_j)$.

Suppose that $s \in I_{i,j}^x$. Then we have $0 < D_i^\diamond f(x + se_j) \leq D_i^\diamond f(x + te_j)$ and so we have $s, t \in I_{i,j}^x$ and $f(x + se_j) \leq f(x + te_j)$. Since ϕ' is increasing and nonnegative, we have $0 \leq \phi'(f(x + se_j)) \leq \phi'(f(x + te_j))$. So by (i) we have $D_i^\diamond h(x + se_j) \leq D_i^\diamond h(x + te_j)$.

Suppose that $t \in J_{i,j}^x$. Then we have $D_i^\diamond f(x + se_j) \leq D_i^\diamond f(x + te_j) < 0$ and so we have $s, t \in J_{i,j}^x$ and $f(x + te_j) \leq f(x + se_j)$. Since ϕ' is increasing and nonnegative, we have $0 \leq \phi'(f(x + te_j)) \leq \phi'(f(x + se_j))$. So by (i) we have $D_i^\diamond h(x + se_j) \leq D_i^\diamond h(x + te_j)$.

Suppose that $s \notin I^x_{i,j}$ and $t \notin J^x_{i,j}$. Then we have $D^\diamond_i f(x+se_j) \le 0 \le D^\diamond_i f(x+te_j)$ and so by (i) and nonnegativity of ϕ', we have $D^\diamond_i h(x+se_j) \le 0 \le D^\diamond_i h(x+te_j)$.

Hence, in all cases we have $D^\diamond_i h(x+se_j) \le D^\diamond_i h(x+te_j)$ and so we see that $D^\diamond_i h(x)$ is increasing in the jth coordinate. So by (2), we have $\Delta_{ij}(\phi \circ f) \ge 0$. Since $\xi : J \to \mathbf{R}$ is increasing and convex, there exist continuously differentiable, increasing convex functions $\xi_1, \xi_2, \ldots : \mathbf{R} \to \mathbf{R}$ such that $\xi_m(t) \to \xi(t)$ for all $t \in J$. By the argument above, we have $\Delta_{ij}(\xi_m \circ f) \ge 0$ for all $m \ge 1$ and since $\xi_m(f(x)) \to \xi(f(x))$, we see that $\Delta_{ij}(\xi \circ f) \ge 0$. Thus (3) is proved and (4) follows in the same manner. □

Corollary 4.5. *Let $\phi_1, \ldots, \phi_k : \mathbf{R} \to \mathbf{R}$ be given functions which are either all increasing or all decreasing and let us define $\phi(x) = (\phi_1(x_1), \ldots, \phi_k(x_k))$ and*

$$M_k(x) = \max_{i \in [k]} x_i \ , \ m_k(x) = \min_{i \in [k]} x_i \ , \ Q_k(x) = \max_{i,j \in [k]} |x_i - x_j|$$

for all $x = (x_1, \ldots, x_k) \in \mathbf{R}^k$. Let $\psi : \mathbf{R} \to \mathbf{R}$ be an increasing function and let $\xi : [0, \infty) \to \mathbf{R}$ be an increasing convex function. Then we have

(1) *$M_k(x)$ and $G(x) := \psi(M_k(\phi(x)))$ are submodular.*

(2) *$m_k(x)$ and $F(x) := \psi(m_k(\phi(x)))$ are supermodular.*

(3) *$Q_k(x)$ and $H(x) := \xi(Q_k(\phi(x)))$ are submodular.*

Proof. Let $x, y \in \mathbf{R}^k$ be given. Then we have $M_k(x \vee y) = M_k(x) \vee M_k(y)$ and since M_k is increasing, we have $M_k(x \wedge y) \le M_k(x) \wedge M_k(y)$.

Hence we have

$$M_k(x \vee y) + M_k(x \wedge y) \le M_k(x) \vee M_k(y) + M_k(x) \wedge M_k(y) = M_k(x) + M_k(y).$$

Hence, we see that (1) follows from Proposition 4.1 and since $m_k(x) = -M_k(-x)$, we see that (2) follows from (1) and Proposition 4.1. Since $Q_k(x) = M_k(x) - m_k(x)$, we see that Q_k is submodular. Let $1 \le i \ne j \le k$ be given and set $\pi_{ij} = [k] \setminus \{i, j\}$. Let $x = (x_1, \ldots, x_k) \in \mathbf{R}^k$ be given and set $M_{ij} = \max_{\nu \in \pi_{ij}} x_\nu$ and $m_{ij} = \min_{\nu \in \pi_{ij}} x_\nu$ with the usual conventions $\sup \emptyset := -\infty$ and $\inf \emptyset := +\infty$. Then we have $\overline{D}^r_i Q_k(x) = 1$ if $x_i \ge x_j \vee M_{ij}$, $\overline{D}^r_i Q_k(x) = 0$ if $x_j \wedge m_{ij} \le x_i < x_j \vee M_{ij}$ and $\overline{D}^r_i Q_k(x) = -1$ if $x_i < x_j \wedge m_{ij}$. So by Theorem 4.3.(4), we have $\Delta_{ij}(\xi \circ Q_k) \le 0$. Hence, we see that (3) follows from (4.3) and Proposition 4.1. □

Lemma 4.6. *Let $D \subseteq \mathbf{R}$ be a countable set and let ℓ_D denote the topology on $\mathbf{R}$ generated by $\{(a, b] \mid a, b \in \mathbf{R}\} \cup \{G \mid G \subseteq D\}$. Let $\Theta(\mathbf{R}^k)$ denote the set of all $\theta : \mathbf{R}^k \to \mathbf{R}^k$ of the form $\theta(x_1, \ldots, x_k) = (\theta_1(x_1), \ldots, \theta_k(x_k))$ for some increasing, right continuous step functions $\theta_1, \ldots, \theta_k : \mathbf{R} \to \mathbf{R}$. Let $\Phi \subseteq \mathbf{R}^{\mathbf{R}^k}$ be a non-empty set and let $f : \mathbf{R}^k \to \mathbf{R}$ be a given function satisfying*

(1) *Φ is sequentially closed and $f \circ \theta \in \Phi \ \ \forall \theta \in \Theta(\mathbf{R}^k)$.*

(2) *f is ℓ_D-continuous in the ith coordinate for all $i = 1, \ldots, k$.*

Then we have $f \in \Phi \cap M(\mathbf{R}^k, \mathcal{B}^k)$. More precisely, $f \in \Phi$ and f is of Baire class $k+1$. Let $F, H \in M(\mathbf{R}^k, \mathcal{B}^k)$ be given functions satisfying

(3) $\exists\delta > 0$ *so that* $F(x) \le f(x-y) \le H(x) \ \ \forall x \in \mathbf{R}^k \ \forall y \in [0,\delta]^k$

and let $\mathfrak{F}$ *denote the set of all locally bounded, right continuous functions* $h : \mathbf{R}^k \to \mathbf{R}$ *such that* $F(x) \le h(x) \le H(x)$ *for all* $x \in \mathbf{R}^k$. *Let* μ *and* ν *be Borel measures on* $\mathbf{R}^k$ *satisfying*

(4) $F \in L^1(\mu)$, $H \in L^1(\nu)$ *and* $\int^* \phi \, d\mu \le \int^* \phi \, d\nu \ \ \forall \phi \in \Phi \cap \mathfrak{F}$.

Then we have $f \in L^1(\mu)$ *and* $\int_{\mathbf{R}^k} f \, d\mu \le \int^* f \, d\nu < \infty$.

Remark. (a): Note that $g : \mathbf{R} \to \mathbf{R}$ is ℓ_D-continuous if and only if g is left continuous at x for all $x \in \mathbf{R} \backslash D$. Recall that $g : \mathbf{R} \to \mathbf{R}$ is a *right continuous step function* if and only if there exist numbers $(c_i \mid i \in \mathbf{Z})$ such that $c_i < c_{i+1}$ for all $i \in \mathbf{Z}$, $\sup_{i\in\mathbf{Z}} c_i = \infty$, $\inf_{i\in\mathbf{Z}} c_i = -\infty$ and $g(t) = g(c_i)$ for all $t \in [c_i, c_{i+1})$ and all $i \in \mathbf{Z}$.

(b): Let T be a topological space and let $h : T \to \mathbf{R}$ be a function. Recall that h is of *Baire class* 0 if h is continuous and that h is of *Baire class* α for some ordinal $\alpha > 0$ if and only if h is a pointwise limit of a sequence of functions of Baire class $< \alpha$.

Proof. Let $E_1 \subseteq E_2 \subseteq \cdots \subseteq \mathbf{R}$ be an increasing sequence of finite sets such that $E_n \uparrow D$ and set $D_n = E_n \cup \{i\,2^{-n} \mid i \in \mathbf{Z}\}$ and $\theta_n(t) = \sup(D_n \cap (-\infty, t])$ for all $n \ge 1$ and all $t \in \mathbf{R}$. Then θ_n is an increasing, right continuous, step function and we have $t - 2^{-n} \le \theta_n(t) \le \theta_{n+1}(t) \le t$ for all $n \ge 1$ and all $t \in \mathbf{R}$. Let $\sigma_1, \ldots, \sigma_k \in \mathbf{N}$ and $x = (x_1, \ldots, x_h) \in \mathbf{R}^k$ be given and set $f^k_{\sigma_1,\ldots,\sigma_k}(x) = f(\theta_{\sigma_1}(x_1), \ldots, \theta_{\sigma_k}(x_k))$ and

$$f^i_{\sigma_1,\ldots,\sigma_i}(x) = f(\theta_{\sigma_1}(x_1), \ldots, \theta_{\sigma_i}(x_i), x_{i+1}, \ldots, x_k) \quad \text{for } 1 \le i < k.$$

Note that $\theta_n(t) = t$ for all $t \in E_m$ and all $n \ge m$. Since $E_n \uparrow D$, we see that $\theta_n(t) \to t$ in ℓ_D and so by (2) we have

$$f^{i-1}_{\sigma_1,\ldots,\sigma_{i-1}}(x) = \lim_{\sigma_i \to \infty} f^i_{\sigma_1,\ldots,\sigma_i}(x) \ \ \forall 1 < i \le k \ , \ f(x) = \lim_{\sigma_1 \to \infty} f^1_{\sigma_1}(x).$$

By (1), we have $f^k_{\sigma_1,\ldots,\sigma_k} \in \Phi$ and since Φ is sequentially closed we see that $f \in \Phi$. Since θ_n is a right continuous step function, we see that $f^k_{\sigma_1,\ldots,\sigma_k}$ is a right continuous, locally bounded Borel function on $\mathbf{R}^k$. So by Lemma 2.2 we see that $f^k_{\sigma_1,\ldots,\sigma_k}$ is of Baire class 1. Hence, we see that f is of Baire class $k+1$.

Let μ and ν be Borel measures on $\mathbf{R}^k$ and let $F \in L^1(\mu)$ and $H \in L^1(\nu)$ be given functions satisfying (3)–(4). Let $\delta > 0$ be chosen according to (3) and let $q \in \mathbf{N}$ be chosen such that $2^{-q} < \delta$. Let Λ denote the set of all functions $h : \mathbf{R}^k \to \mathbf{R}$ satisfying $\int^* h \, d\mu \le \int^* h \, d\nu$ and set $\Psi = \{\psi \in \Phi \mid F \le \phi \le H\}$. By (1), we see that Ψ is sequentially closed and since $F \in L^1(\mu)$ and $H \in L^1(\nu)$, we see that $\{\psi^- \mid \psi \in \Psi\}$ is uniformly μ-integrable and that $\{\psi^+ \mid \psi \in \Psi\}$ is uniformly ν-integrable. By Theorem 3.3.(9), we have that $\Lambda \cap \Psi \cap M(\mathbf{R}^k, \mathcal{B}^k)$ is sequentially closed. Let $\sigma_1, \ldots, \sigma_k \ge q$ be given integers. Since θ_n is a right continuous, step function satisfying $t - \delta \le \theta_n(t) \le t$ for all $n \ge q$ and all $t \in \mathbf{R}$, we have $f^k_{\sigma_1,\ldots,\sigma_k} \in \Phi \cap \mathfrak{F} \cap M(\mathbf{R}^k, \mathcal{B}^k)$. So by (3)–(4), we have $f^k_{\sigma_1,\ldots,\sigma_k} \in \Lambda \cap \Psi \cap M(\mathbf{R}^k, \mathcal{B}^k)$

for all $\sigma_1, \dots, \sigma_k \geq q$ and since $\Lambda \cap \Psi \cap M(\mathbf{R}^k, \mathcal{B}^k)$ is sequentially closed, we see that $f \in \Lambda \cap \Psi \cap M(\mathbf{R}^k, \mathcal{B}^k)$. Hence, we have

$$-\infty < \int_{\mathbf{R}^k} F\, d\mu \leq \int_* f\, d\mu \leq \int^* f\, d\mu \leq \int^* f\, d\nu \leq \int_{\mathbf{R}^k} H\, d\nu < \infty.$$

So by Theorem 2.1.(8) in [4] we have $f \in L^1(\mu)$ and $\int_{\mathbf{R}^k} f\, d\mu \leq \int^* f\, d\nu < \infty$ $\square$

Theorem 4.7. *Let μ and ν be Borel probability measures on $\mathbf{R}^k$ with one-dimensional marginals $\mu_1, \dots, \mu_k$ and $\nu_1, \dots, \nu_k$, respectively. Then we have*

(1) $\mu \preceq_{\mathrm{bm}} \nu \Leftrightarrow \mu_i = \nu_i$ *for* $i = 1, \dots, k$.

(2) $\displaystyle\int_{\mathbf{R}^k} f\, d\mu \leq \int_{\mathbf{R}^k} f\, d\nu \ \forall f \in C_b^\infty(\mathbf{R}^k) \cap \mathrm{ism}(\mathbf{R}^k) \ \Leftrightarrow \mu \preceq_{\mathrm{ism}} \nu$.

(3) $\displaystyle\int_{\mathbf{R}^k} f\, d\mu \leq \int_{\mathbf{R}^k} f\, d\nu \ \forall f \in C_b^\infty(\mathbf{R}^k) \cap \mathrm{m}(\mathbf{R}^k) \ \Leftrightarrow \mu \preceq_{\mathrm{bm}} \nu$.

(4) $\mu \preceq_{\mathrm{bsm}} \nu \Leftrightarrow \mu \preceq_{\mathrm{ism}} \nu$ *and* $\mu \preceq_{\mathrm{bm}} \nu$.

(5) $\displaystyle\int_{\mathbf{R}^k} f\, d\mu \leq \int_{\mathbf{R}^k} f\, d\nu \ \forall f \in C_b^\infty(\mathbf{R}^k) \cap \mathrm{sm}(\mathbf{R}^k) \ \Leftrightarrow \mu \preceq_{\mathrm{bsm}} \nu$.

Proof. Throughout the proof we let Λ denote the set of all functions $f : \mathbf{R}^k \to \mathbf{R}$ satisfying $\int^* f\, d\mu \leq \int^* f\, d\nu$. (1) follows easily from Proposition 4.2.(4).

(2): Suppose that $C_b^\infty(\mathbf{R}^k) \cap \mathrm{ism}(\mathbf{R}^k) \subseteq \Lambda$. Let C_b^r denote the set of all bounded, right continuous functions on $\mathbf{R}^k$ and set $\kappa(x) \equiv 1$. Then $(\mathrm{ism}(\mathbf{R}^k), \kappa)$ satisfies conditions (1)–(2) in Theorem 3.4 and so we conclude that $C_b^r \cap \mathrm{ism}(\mathbf{R}^k) \subseteq \Lambda$. Let $f \in B(\mathbf{R}^k) \cap \mathrm{ism}(\mathbf{R}^k)$ be given. Then there exists $c > 0$ such that $|f(x)| \leq c$ for all $x \in \mathbf{R}^k$. Let $A \subseteq \mathbf{R}^k$ be a countable dense subset of $\mathbf{R}^k$. Since f_i^u is increasing, there exists a countable set $D \subseteq \mathbf{R}$ such that $\mathbf{R} \setminus D \subseteq C(f_i^u)$ for all $i \in [k]$ and all $u \in A$. Hence, by Proposition 4.2.(3), we have $\mathbf{R} \setminus D \subseteq C(f_i^u)$ for all $i \in [k]$ and all $u \in \mathbf{R}^k$ and so we see the (f, D) satisfies condition (2) in Lemma 4.6. By Proposition 4.1.(2) we see that $(\mathrm{ism}(\mathbf{R}^k, f)$ satisfies condition (1) in Lemma 4.6. Since $C_b^r \cap \mathrm{ism}(\mathbf{R}^k) \subseteq \Lambda$ we see that f satisfies condition (3)–(4) in Lemma 4.6 with $F(x) \equiv -c$ and $H(x) \equiv c$. Hence, we have $f \in \Lambda$ for all $f \in B(\mathbf{R}^k) \cap \mathrm{ism}(\mathbf{R}^k)$. Let $f \in \mathrm{ism}(\mathbf{R}^k)$ be given. Set $b^m = (m, \dots, m)$ and $f_{m,n}(x) = f(x \wedge b^m) \vee (-n)$ for all $m, n \in \mathbf{N}$ and all $x \in \mathbf{R}^k$. Then $f_{m,n}$ is increasing and bounded and by Proposition 4.1, we have that $f_{m,n}$ is supermodular. Hence, we have $f_{m,n} \in \Lambda$. Since f is increasing, we see that $-n \leq f_{1,n} \leq f_{2.n} \leq \cdots$ and $\sup_{m \geq 1} f_{m,n}(x) = f(x) \vee (-n)$ for all $x \in \mathbf{R}^k$. So by Theorem 3.3.(8) we see that $f(x) \vee (-n) \in \Lambda$ for all $n \in \mathbf{N}$ and so by Theorem 3.3.(2) we have $f \in \Lambda$; that is, $\mu \preceq_{\mathrm{ism}} \nu$ which completes the proof of (2).

(3): Suppose that $C_b^\infty(\mathbf{R}^k) \cap \mathrm{m}(\mathbf{R}^k) \subseteq \Lambda$ and let $i \in [k]$ be given. By Proposition 4.2.(4), we have $\int_{\mathbf{R}} \phi\, \mu_i = \int_{\mathbf{R}} \phi\, \nu_i$ for all $\phi \in C_b^\infty(\mathbf{R})$ and so we have $\mu_i = \nu_i$. Hence, we see that (3) follows from (1).

(4): The implication "$\Rightarrow$" in (4) follows directly from (2). Suppose that $\mu \preceq_{\mathrm{ism}} \nu$ and $\mu \preceq_{\mathrm{bm}} \nu$ and let $f \in \mathrm{bsm}(\mathbf{R}^k)$ be given. Set $a^n = (-n, \dots, -n)$ for $n \geq 1$.

By Proposition 4.2.(1) we have that the limits

$$\alpha_i(t) := \lim_{n\to\infty} (f_i^{a^n}(t) - f_i^{a^n}(0))$$

exist and are finite for all $t \in \mathbf{R}$ and all $i \in [k]$. Since f is a bounded Borel function, we see that α_i is bounded Borel function. So by Proposition 4.2.(4), we have that $G(x) := \sum_{i=1}^k \alpha_i(x_i)$ is a bounded, modular, Borel function. By Proposition 4.2.(5), we see that

$$x \curvearrowright f(x \vee a) - \sum_{i=1}^k (f_i^a(x_i \vee a_i) - f_i^a(0))$$

is increasing and supermodular for all $a \in \mathbf{R}^k$. Hence, we see that $F(x) := f(x) - G(x)$ is a bounded, increasing, supermodular, Borel function. Since $\mu \preceq_{\text{ism}} \nu$, we have $F \in \Lambda \cap L^1(\mu) \cap L^1(\nu)$ and since $\mu \preceq_{\text{bm}} \nu$, we have $G \in \Lambda \cap L^1(\mu) \cap L^1(\nu)$. So by Theorem 3.3.(1) we have $f = F + G \in \Lambda$ for all $f \in \text{bsm}(\mathbf{R}^k)$ which completes the proof of (4).

(5): Suppose that $f \in C_b^\infty(\mathbf{R}^k) \cap \text{sm}(\mathbf{R}^k) \subseteq \Lambda$. By (2) and (3), we have $\mu \preceq_{\text{bm}} \nu$ and $\mu \preceq_{\text{ism}} \nu$. So by (4) we have $\mu \preceq_{\text{bsm}} \nu$. □

Theorem 4.8. *Let μ and ν be Borel probability measures on $\mathbf{R}^k$ such that $\mu \preceq_{\text{bsm}} \nu$ and let $\mu_1, \ldots, \mu_k$ denote the one-dimensional marginals of μ. Let $f : \mathbf{R}^k \to \mathbf{R}$ be a supermodular Borel function and let us define $f_{\vee c}(x) = f(x \vee c)$ and $f_{\wedge c}(x) = f(x \wedge c)$ for all $c, x \in \mathbf{R}^k$. If $c \in \mathbf{R}^k$ is a given vector, then we have*

(1) *If f is either increasing or decreasing, then $\int^* f\, d\mu \le \int^* f\, d\nu$.*

(2) *If $f_i^c \in L^1(\mu_i)$ for all $i \in [k]$, then we have $\int^* f_{\vee c} d\mu \le \int^* f_{\vee c}\, d\nu$ and $\int^* f_{\wedge c} d\mu \le \int^* f_{\wedge c}\, d\nu$.*

Let $A, B \subseteq \mathbf{R}^k$ be given sets satisfying $\cup_{a\in A}[a, *] = \mathbf{R}^k = \cup_{b\in B}[*, b]$. Then we have $\int^* f\, d\mu \le \int^* f\, d\nu$ if just one of the following three conditions holds:

(A) *$f_i^a \in L^1(\mu_i)\ \forall i \in [k]\ \forall a \in A$, $\{f_{\vee a}^- \mid a \in A\}$ is uniformly μ-integrable and $\{f_{\vee a}^+ \mid a \in A\}$ is uniformly ν-integrable.*

(B) *$f_i^b \in L^1(\mu_i)\ \forall i \in [k]\ \forall b \in B$, $\{f_{\wedge b}^- \mid b \in B\}$ is uniformly μ-integrable and $\{f_{\wedge b}^+ \mid b \in B\}$ is uniformly ν-integrable.*

(C) *There exist functions $h_1 \in L_+^1(\mu_1), \ldots, h_k \in L_+^1(\mu_k)$ such that $|f(x)| \le \sum_{i=1}^k h_i(x_i)$ for all $x = (x_1, \ldots, x_k) \in \mathbf{R}^k$.*

Proof. Throughout the proof, we let Λ denote the set of all functions $h : \mathbf{R}^k \to \mathbf{R}$ satisfying $\int^* h\, d\mu \le \int^* h\, d\nu$. We set $\tilde{f}(x) = f(-x)$ for all $x \in \mathbf{R}^k$ and we set $\tilde{\mu}(B) = \mu(-B)$ and $\tilde{\nu}(B) - \nu(-B)$ for all $B \in \mathcal{B}^k$. Then $\tilde{\mu}$ and $\tilde{\nu}$ are Borel probability measures on $\mathbf{R}^k$ such that $\int \tilde{f}\, d\tilde{\mu} = \int f\, d\mu$ and $\int \tilde{f}\, d\tilde{\nu} = \int f\, d\nu$. By Proposition 4.1, we see that $\tilde{f}$ is supermodular and that $\tilde{\mu} \preceq_{\text{bsm}} \tilde{\nu}$.

(1): Since $\mu \preceq_{\text{bsm}} \nu$, we have $\text{bsm}(\mathbf{R}^k) \subseteq \Lambda$ and by Theorem 4.7, we have $\text{ism}(\mathbf{R}^k) \subseteq \Lambda$ and $\mu_i = \nu_i$ for $i = 1, \ldots, k$ where $\nu_1, \ldots, \nu_k$ are the one-dimensional marginals of ν. Hence, we have $f \in \Lambda$ if f is increasing. Applying this to the triple $(\tilde{f}, \tilde{\mu}, \tilde{\nu})$, we see that $f \in \Lambda$ if f is decreasing.

(2): Suppose that $f_i^c \in L^1(\mu_i)$ for all $i = 1, \ldots, k$ and set $G(x) = \sum_{i=1}^k f_i^c(x_i \vee c_i)$ for all $x = (x_1, \ldots, x_k) \in \mathbf{R}^k$. By Proposition 4.2.(4), we have that G is a modular, Borel function and since μ_i is finite, we have that $f_i^c(t \vee c_i)$ belongs to $L^1(\mu_i)$. Since $\nu_i = \mu_i$, we have $G \in \Lambda \cap L^1(\mu) \cap L^1(\nu)$ and by Proposition 4.2.(5), we have $f_{\vee c} - G \in \mathrm{ism}(\mathbf{R}^k) \subseteq \Lambda$. So by Theorem 3.3.(1) we have $f_{\vee c} \in \Lambda$. Applying this to triple $(\tilde{f}, \tilde{\mu}, \tilde{\nu})$ with $c := -c$, we see that $f_{\wedge c} \in \Lambda$.

Suppose that condition (A) holds. By (2), we see that $f_{\vee a} \in \Lambda$ for all $a \in A$ and since $\mathbf{R}^k = \cup_{a \in A} [a, *]$, there exists $a^n = (a_1^n, \ldots, a_k^n) \in A$ such that $a_i^{n+1} \leq a_i^n \leq -n$ for all $n \in \mathbf{N}$ and all $i \in [k]$. Then we have $f_{\vee a^n}(x) \to f(x)$ and by (2) and (A), we have $f_{\vee a^n} \in \Lambda$, $\{f_{\vee a^n}^- \mid n \geq 1\}$ is uniformly μ-integrable and $\{f_{\vee a^n}^+ \mid n \geq 1\}$ is uniformly ν-integrable. So by (3.3)–(3.4) and Theorem 3.3.(5) we have $f \in \Lambda$.

Suppose that condition (B) holds. Applying case (A) on the triple $(\tilde{f}, \tilde{\mu}, \tilde{\nu})$ with $A := -B$, we see that $f \in \Lambda$. Suppose that condition (C) holds. Since $h_i \in L_+^1(\mu_i)$, we see that $f_i^x \in L^1(\mu_i)$ for all $x \in \mathbf{R}^k$ and all $i \in [k]$. Let ξ be a finite Borel measure on $\mathbf{R}$. Then I claim that we have

(i) $\displaystyle \liminf_{u \to -\infty} \int_{\mathbf{R}} |h(t) - h(t \vee u)| \xi(dt) = 0 \quad \forall h \in L^1(\xi).$

Proof of (i). Suppose that (i) fails. Then there exist $h \in L^1(\xi)$ and a positive number $\delta > 0$ such that $\liminf_{u \downarrow -\infty} \int_{\mathbf{R}} |h(t) - h(t \vee u)| \, \xi(dt) > 2\delta$. Since $h \in L^1(\xi)$, there exists $q \in \mathbf{R}$ such that

$$\int_{(-\infty, q]} |h| \, d\xi < \delta \text{ and } \int_{\mathbf{R}} |h(t) - h(t \vee u)| \, \xi(dt) > 2 \, \delta \quad \forall u \leq q.$$

Let $u \leq q$ be given and set $F_\xi(u) := \xi((-\infty, u])$. Then we have

$$\begin{aligned} 2\delta &< \int_{\mathbf{R}} |h(t) - h(t \vee u)| \, \xi(dt) \leq \int_{(-\infty, q]} |h(t)| \xi(dt) + |h(u)| F_\xi(u) \\ &\leq \delta + |h(u)| F_\xi(u) \end{aligned}$$

and so we see that $|h(u)| \, F_\xi(u) > \delta$ for all $u \leq q$. Set $m = \inf_{s \leq q} |h(s)|$ and let $s \leq q$ be given. Since F_ξ is increasing. we have $\delta \leq |h(s)| \, F_\xi(s) \leq |h(s)| \, F_\xi(q)$ and so we have $\delta \leq m \, F_\xi(q)$ and

$$\delta \leq m F_\xi(q) = \textstyle\int_{(-\infty, q]} m \, \xi(ds) \leq \displaystyle\int_{(-\infty, q]} |h(s)| \, \xi(ds) < \delta$$

which is impossible. Thus, we see that (i) holds.

Let $i \in [k]$ be given. By (i) there exist numbers $a_i^1 > a_i^2 > \cdots$ such that $a_i^n < -n$ for all $n \geq 1$ and

$$\lim_{n \to \infty} \int_{\mathbf{R}} |h_i(t) - h_i(t \vee a_i^n)| \, \mu_i(dt) = 0 \quad \forall i = 1, \ldots, k.$$

Set $a^n = (a_1^n, \dots, a_k^n)$ and $H(x) = \sum_{i=1}^k h_i(x_i)$ for $x = (x_1, \dots, x_k) \in \mathbf{R}^k$. Since $\mu_i = \nu_i$ and

$$|H(x) - H(x \vee a^n)| \leq \sum_{i=1}^n |h_i(x_i) - h(x_i \vee a_i^n)|$$

we have

$$\int_{\mathbf{R}^k} |H(x) - H(x \vee a^n)| \, \mu(dx) \to 0 \ , \ \int_{\mathbf{R}^k} |H(x) - H(x \vee a^n)| \, \nu(dx) \to 0.$$

In particular, we see that $\{H_{\vee a^n} \mid n \geq 1\}$ is uniformly μ-integrable and uniformly ν-integrable. Since $|f(x \vee a^n)| \leq H(x \vee a^n)$, we see that $\{f_{\vee a^n} \mid n \geq 1\}$ is uniformly μ-integrable and uniformly ν-integrable. So by case (A) we have $f \in \Lambda$ $\square$

References

[1] Dunford, N. and Schwartz, J.T. (1950): *Linear Operator, Part 1.* Interscience Publishers, Inc., New York.

[2] Fernique, X. (2012): *Regularité des trajectoires des functions alétoires Gaussiennes.* Probability in Banach Spaces at Saint Flour (Ecole d'Eté de Probabilités de Saint Flour IV, 1974), Springer Verlag, Berlin and Heidelberg.

[3] Hoffmann-Jørgensen, J. (1994): *Probability with a view towards statistics*, vol. 1+2. Chapman & Hall, New York and London.

[4] Hoffmann-Jørgensen, J. (2006): *Constructions of contents and measures satisfying a prescribed set of integral inequalities.* Probab. Math. Stat. 26, pp. 283–313.

[5] Horn, A.R and Johnson, C.R. (1990): *Matrix analysis.* Cambridge University Press, London and New York.

[6] Joag-Dev, K., Perlman, M.D. and Pitt, L.D. (1983): *Association of normal random variables and Slepian's inequality.* Ann. Probab. 11, pp. 451–455.

[7] Kahane, J.-P. (1986): *Une inégalité du type Slepian et Gordon sur les processus gaussienns.* Israel J. Math. 55, pp. 109–110.

[8] Kemperman, J.H.B. (1977): *On the FKG-inequality for measures on a partially ordered space.* Indag. Math. 39, pp. 313–331.

[9] Ledoux, M. and Talagrand, M. (1991): *Probability in Banach spaces.* Springer Verlag, Berlin and New York.

[10] Müller, A. and Stoyan, D. (2002): *Comparison methods for stochastic models and risks.* John Wiley & Sons, Ltd., Chichester England.

[11] Ostrowski, A. (1929): *Mathematische Mizellen XIV über die Funktionalgleichung der Exponentialfunktion und verwandte Funktionalgleichungen.* Jahreber. Deutsch. Math.-Verein. 38, pp. 54–62.

[12] Pitt, L.D. (1982): *Positively correlated normal variables are associated.* Ann. Probab. 10, pp. 496–499.

[13] Robertson, A.P. and Robertson, W. (1964): *Topological vector spaces.* Cambridge University Press, Cambridge.

[14] Rudin, W. (1962): *Fourier analysis on groups.* Interscience Publishers (a division of John Wiley & Sons), New York and London.

[15] Saks, S. (1937): *Theory of the integral.* Hafner Publishing Company, New York.

[16] Samorodnitsky, G. and Taqqu, M.S. (1993): *Stochastic monotonicity and Slepian-type inequalities for infinitely divisible and stable random vectors.* Ann. Probab. 21, pp. 143–160.

[17] Simons, G. (1977): *An unexpected expectation.* Ann.Probab. 5, pp. 157–158.

[18] Slepian, D. (1962): *The one-sided barrier problem for Gaussian noise.* Bell Syst. Tech. J. 41, pp. 463–501.

[19] Tchen, A.H. (1980): *Inequalities for distributions with given marginals.* Ann. Probab. 8, pp. 814–827.

[20] Wright, J.D.M. (1975): *The continuity of mid-point convex functions.* Bull. London Math. Soc. 7, pp. 89–92.

J. Hoffmann-Jørgensen
Department of Mathematics
University of Aarhus
Denmark
e-mail: hoff@imf.au.dk