Fisher information and matrix-valued valuations

Monika Ludwig *

Abstract

All affinely contravariant matrix-valued valuations on the Sobolev space $W^{1,2}(\mathbb{R}^n)$ are completely classified. It is shown that there is a unique such valuation. This valuation turns out to be the Fisher information matrix.

2000 AMS subject classification: 62B10 (46E35, 52A20, 52B45, 94A17).

A function Z defined on a lattice $(\mathcal{L}, \vee, \wedge)$ and taking values in an abelian semigroup is called a valuation if

$$\mathrm{Z}(f \vee g) + \mathrm{Z}(f \wedge g) = \mathrm{Z}(f) + \mathrm{Z}(g) \tag{1}$$

for all $f, g \in \mathcal{L}$. A function Z defined on some subset $\mathcal{S}$ of $\mathcal{L}$ is called a valuation on $\mathcal{S}$ if (1) holds whenever $f, g, f \vee g, f \wedge g \in \mathcal{S}$. Results on valuations on compact convex sets in $\mathbb{R}^n$ are classical and start with Dehn's solution of Hilbert's Third Problem in 1901. See [24, 27] for information on the classical theory of valuations on convex sets and [1–5, 9, 11, 16, 19–22, 28–31, 33, 34, 43–45] for some of the more recent results. Valuations were also investigated on star shaped sets [25, 26] and on manifolds [6–8, 10].

In this paper, we classify matrix-valued valuations on the Sobolev space $W^{1,2}(\mathbb{R}^n)$, that is, the space of functions belonging to $L^2(\mathbb{R}^n)$ whose distributional first-order derivatives belong to $L^2(\mathbb{R}^n)$. The lattice $(W^{1,2}(\mathbb{R}^n), \vee, \wedge)$ is defined by letting $f \vee g$ denote the maximum and $f \wedge g$ the minimum of $f, g \in W^{1,2}(\mathbb{R}^n)$. Let $\langle \mathbb{M}^n, + \rangle$ be the additive group of real symmetric $n \times n$ matrices. As in the classical results for valuations on convex sets, we use invariance and covariance properties with respect to suitable transformation groups to classify valuations. Since we are interested in operators

*Research supported, in part, by NSF grant DMS-0805623

that do not depend on the choice of the coordinate system, we use the general linear group $\mathrm{GL}(n)$. An operator $\mathrm{Z} : W^{1,2}(\mathbb{R}^n) \to \mathbb{M}^n$ is called $\mathrm{GL}(n)$ contravariant if for some $p \in \mathbb{R}$,

$$\mathrm{Z}(f \circ \phi^{-1}) = |\det \phi|^p \; \phi^{-t}\,\mathrm{Z}(f)\,\phi^{-1}$$

for all $f \in W^{1,2}(\mathbb{R}^n)$ and $\phi \in \mathrm{GL}(n)$, where $\det \phi$ is the determinant of ϕ and ϕ^{-t} denotes the inverse of the transpose of ϕ. It is called translation invariant if $\mathrm{Z}(f \circ \tau^{-1}) = \mathrm{Z}(f)$ for all $f \in W^{1,2}(\mathbb{R}^n)$ and translations τ, and homogeneous if for some $q \in \mathbb{R}$, we have $\mathrm{Z}(sf) = |s|^q\,\mathrm{Z}(f)$ for all $f \in W^{1,2}(\mathbb{R}^n)$ and $s \in \mathbb{R}$. An operator $\mathrm{Z} : W^{1,2}(\mathbb{R}^n) \to \mathbb{M}^n$ is called affinely contravariant if it is $\mathrm{GL}(n)$ contravariant, translation invariant and homogeneous.

Theorem. *An operator $\mathrm{Z} : W^{1,2}(\mathbb{R}^n) \to \langle\, \mathbb{M}^n, + \,\rangle$, where $n > 2$, is a continuous and affinely contravariant valuation if and only if there is a constant $c \in \mathbb{R}$ such that*

$$\mathrm{Z}(f) = c\,\mathrm{J}(f^2)$$

for every $f \in W^{1,2}(\mathbb{R}^n)$.

Here $\mathrm{J}(g)$ is the Fisher information matrix of a weakly differentiable function $g : \mathbb{R}^n \to [0, \infty)$, that is, the $n \times n$ matrix with entries

$$\mathrm{J}_{ij}(g) = \int_{\mathbb{R}^n} \frac{\partial \log g(x)}{\partial x_i}\,\frac{\partial \log g(x)}{\partial x_j}\,g(x)\,dx. \tag{2}$$

The Fisher information matrix plays an important role in statistics and information theory (see [14,15]). In general, Fisher information is a measure of the minimum error in the maximum likelihood estimate of a parameter in a distribution. The Fisher information matrix (2) describes such an error for a random vector of density g with respect to a location parameter.

The proof of the theorem is based on the intriguing connection between information theory and the L^2 Brunn Minkowski Theory (see [12, 13, 17, 23, 35–42, 47, 48] for information on the L^p Brunn Minkowski Theory). Lutwak, Yang, and Zhang [37] introduced a new ellipsoid associated with convex sets and showed in [39] that this LYZ ellipsoid corresponds to the Fisher information ellipsoid defined by the Fisher information matrix. In [29], it was shown that the matrix corresponding to the LYZ ellipsoid is the only matrix-valued valuations on convex sets that is $\mathrm{GL}(n)$ contravariant with $p \geq 0$. The proof of the theorem makes essential use of this classification result.

1 Background material on convex polytopes

We work in Euclidean n-space, $\mathbb{R}^n$, and we assume that $n > 2$. We denote by $e_1, \dots, e_n$ the vectors of the standard basis of $\mathbb{R}^n$ and write $x = (x_1, \dots, x_n)$ for $x \in \mathbb{R}^n$. Let $x \cdot y = x_1 y_1 + \dots + x_n y_n$ denote the scalar product of $x, y \in \mathbb{R}^n$ and $|x| = \sqrt{x \cdot x}$ the Euclidean norm of $x \in \mathbb{R}^n$. Let $\mathcal{P}^n$ denote the space of compact convex polytopes in $\mathbb{R}^n$ and $\mathcal{P}_0^n$ the subspace of polytopes containing the origin in their interiors. Both spaces are equipped with the usual topology coming from the Hausdorff metric.

The proof of the theorem makes essential use of a classification result of matrix-valued valuations established in [29]. To state the result, we need the following definitions. For $P \in \mathcal{P}_0^n$, the Lutwak-Yang-Zhang matrix, $\mathrm{L}(P)$, of P is defined in [37] by

$$\mathrm{L}_{ij}(P) = \sum_u \frac{a(P,u)}{h(P,u)}\, u_i\, u_j \tag{3}$$

where we sum over all unit normals u of facets of P and where $a(P,u)$ is the $(n-1)$-dimensional volume of the facet with normal u and $h(P,u)$ is the distance from the origin of the hyperplane containing this facet. An operator $\mathrm{Y} : \mathcal{P}_0^n \to \mathbb{M}^n$ is called $\mathrm{GL}(n)$ contravariant of weight $p \in \mathbb{R}$, if

$$\mathrm{Y}(\phi P) = |\det \phi|^p\ \phi^{-t}\, \mathrm{Y}(P)\, \phi^{-1}$$

for all $P \in \mathcal{P}_0^n$ and $\phi \in \mathrm{GL}(n)$. The following result is a special case of Theorem 2 in [29].

Theorem 1. *An operator $\mathrm{Y} : \mathcal{P}_0^n \to \langle \mathbb{M}^n, + \rangle$, where $n > 2$, is a Borel measurable $\mathrm{GL}(n)$ contravariant valuation of weight $p \geq 0$ if and only if there is a constant $c \in \mathbb{R}$ such that*

$$\mathrm{Y}(P) = c\, \mathrm{L}(P)$$

for every $P \in \mathcal{P}_0^n$.

For $n = 2$, there are additional matrix-valued valuations (see [29]).

2 Background material on $W^{1,2}(\mathbb{R}^n)$

For a measurable function $f : \mathbb{R}^n \to \mathbb{R}$, set $\|f\|_2^2 = \int_{\mathbb{R}^n} |f(x)|^2 dx$ and let $L^2(\mathbb{R}^n)$ denote the space of measurable functions f such that $\|f\|_2 < \infty$. Note that a function f belongs to $W^{1,2}(\mathbb{R}^n)$ if and only if $f \in L^2(\mathbb{R}^n)$ and there exists a measurable vector field $\nabla f : \mathbb{R}^n \to \mathbb{R}^n$ such that $|\nabla f| \in L^2(\mathbb{R}^n)$ and

$$\int_{\mathbb{R}^n} \nu(x) \cdot \nabla f(x)\, dx = -\int_{\mathbb{R}^n} f(x) \nabla \cdot \nu(x)\, dx$$

for every compactly supported smooth vector field $\nu : \mathbb{R}^n \to \mathbb{R}^n$. The vector $\nabla f(x)$ is called the weak gradient of f at x. Note that for $\Phi(x) = \phi x + y$ where $\phi \in \mathrm{GL}(n)$ and $y \in \mathbb{R}^n$, we have

$$\nabla(f \circ \Phi^{-1})(x) = \phi^{-t}\, \nabla f(\Phi^{-1}x) \tag{4}$$

almost everywhere for $f \in W^{1,2}(\mathbb{R}^n)$. We say that a sequence $f_k \in W^{1,2}(\mathbb{R}^n)$ converges to $f \in W^{1,2}(\mathbb{R}^n)$ as $k \to \infty$ in $W^{1,2}(\mathbb{R}^n)$ if $\|f_k - f\|_2 \to 0$ and $\|\nabla(f_k - f)\|_2 \to 0$ as $k \to \infty$, where $\|v\|_2^2 = \int_{\mathbb{R}^n} |v(x)|^2 dx$ for vector fields $v : \mathbb{R}^n \to \mathbb{R}^n$.

An operator $\mathrm{Z} : W^{1,2}(\mathbb{R}^n) \to \mathbb{M}^n$ is called $\mathrm{GL}(n)$ contravariant of weight $p \in \mathbb{R}$, if

$$\mathrm{Z}(f \circ \phi^{-1}) = |\det \phi|^p\, \phi^{-t}\, \mathrm{Z}(f)\, \phi^{-1}$$

for all $f \in W^{1,2}(\mathbb{R}^n)$ and $\phi \in \mathrm{GL}(n)$. It is called homogeneous of degree $q \in \mathbb{R}$, if $\mathrm{Z}(sf) = |s|^q\, \mathrm{Z}(f)$ for all $f \in W^{1,2}(\mathbb{R}^n)$ and $s \in \mathbb{R}$. Taking $s = 0$ shows that if $\mathrm{Z} : W^{1,2}(\mathbb{R}^n) \to \mathbb{M}^n$ is homogeneous, then $q \geq 0$. If $q = 0$, then the continuity of Z implies that $\lim_{s\to 0} \mathrm{Z}(sf) = \mathrm{Z}(0) = \mathrm{Z}(f)$. If Z is in addition $\mathrm{GL}(n)$ contravariant, this implies that $\mathrm{Z}(f)$ is the zero matrix for all $f \in W^{1,2}(\mathbb{R}^n)$. Since clearly $\mathrm{Z}(0) = 0$ for Z homogeneous of degree $q > 0$, we obtain for every homogeneous, $\mathrm{GL}(n)$ contravariant $\mathrm{Z} : W^{1,2}(\mathbb{R}^n) \to \mathbb{M}^n$,

$$\mathrm{Z}(0) = 0. \tag{5}$$

For $f, g \in W^{1,2}(\mathbb{R}^n)$, we have $f \vee g, f \wedge g \in W^{1,2}(\mathbb{R}^n)$ and for $x \in \mathbb{R}^n$ a.e.,

$$\nabla(f \vee g)(x) = \begin{cases} \nabla f(x) & \text{when } f(x) > g(x) \\ \nabla g(x) & \text{when } f(x) < g(x) \\ \nabla f(x) = \nabla g(x) & \text{when } f(x) = g(x) \end{cases} \tag{6}$$

and

$$\nabla(f \wedge g) = \begin{cases} \nabla f(x) & \text{when } f(x) < g(x) \\ \nabla g(x) & \text{when } f(x) > g(x) \\ \nabla f(x) = \nabla g(x) & \text{when } f(x) = g(x). \end{cases} \tag{7}$$

Hence $(W^{1,2}(\mathbb{R}^n), \vee, \wedge)$ is a lattice.

Let $L^{1,2}(\mathbb{R}^n) \subset W^{1,2}(\mathbb{R}^n)$ denote the space of piecewise affine functions on $\mathbb{R}^n$, where a function $\ell : \mathbb{R}^n \to \mathbb{R}$ is piecewise affine, if ℓ is continuous and there are finitely many convex polytopes $P_1, \dots, P_m \subset \mathbb{R}^n$ with pairwise disjoint interiors such that the restriction of ℓ to each P_i is affine and $\ell = 0$ outside $P_1 \cup \dots \cup P_m$. Note that the weak partial derivatives of $\ell \in L^{1,2}(\mathbb{R}^n)$ and the pointwise partial derivatives of ℓ are the same almost everywhere. Also note that piecewise affine functions lie dense in $W^{1,2}(\mathbb{R}^n)$.

For $P \in \mathcal{P}_0^n$, define the piecewise affine function ℓ_P by requiring that $\ell_P(0) = 1$, that $\ell_P(x) = 0$ for $x \notin P$, and that ℓ_P is affine on each pyramid with apex at the origin and base equal to a facet of P. Define $P^{1,2}(\mathbb{R}^n) \subset L^{1,2}(\mathbb{R}^n)$ as the set of all ℓ_P with $P \in \mathcal{P}_0^n$. Note that

$$\ell_{\phi P} = \ell_P \circ \phi^{-1} \tag{8}$$

for $\phi \in \mathrm{GL}(n)$. We remark that multiples and translates of $\ell_P \in P^{1,2}(\mathbb{R}^n)$ correspond to linear elements within the theory of finite elements.

3 The operator $f \mapsto \mathrm{J}(f^2)$

In the following lemma, we prove some well known properties of the operator $f \mapsto \mathrm{J}(f^2)$.

Lemma 2. *The operator* $\mathrm{Z} : W^{1,2}(\mathbb{R}^n) \to \langle \mathbb{M}^n, + \rangle$, *defined by* $\mathrm{Z}(f) = c\,\mathrm{J}(f^2)$ *with* $c \in \mathbb{R}$, *is a continuous and affinely contravariant valuation.*

Proof. It follows from (2) that

$$\mathrm{J}_{ij}(f^2) = 4 \int_{\mathbb{R}^n} \frac{\partial f(x)}{\partial x_i} \, \frac{\partial f(x)}{\partial x_j} \, dx. \tag{9}$$

By the Cauchy-Schwarz inequality

$$\Big| \int\limits_{\mathbb{R}^n} \frac{\partial f(x)}{\partial x_i} \, \frac{\partial f(x)}{\partial x_j} \, dx \Big| \le \Big(\int\limits_{\mathbb{R}^n} \Big| \frac{\partial f(x)}{\partial x_i} \Big|^2 \, dx) \Big)^{\frac{1}{2}} \Big(\int\limits_{\mathbb{R}^n} \Big| \frac{\partial f(x)}{\partial x_j} \Big|^2 \, dx \Big)^{\frac{1}{2}} \le \|\nabla f\|_2^2.$$

Thus $\mathrm{J}_{ij}(f^2) < \infty$ for $f \in W^{1,2}(\mathbb{R}^n)$. Equations (6) and (7) imply that $f \mapsto c\,\mathrm{J}(f^2)$ is a valuation. Suppose that $f_k \to f$ in $W^{1,2}(\mathbb{R}^n)$. Since the Cauchy-Schwarz inequality implies that

$$|\,\mathrm{J}_{ij}(f_k^2) - \mathrm{J}_{ij}(f^2)| \le \|\nabla f_k\|_2 \, \|\nabla (f_k - f)\|_2 + \|\nabla f\|_2 \, \|\nabla (f_k - f)\|_2,$$

the operator $f \mapsto c\,\mathrm{J}(f^2)$ is continuous on $W^{1,2}(\mathbb{R}^n)$.

By (9),

$$\mathrm{J}(f^2) = 4 \int_{\mathbb{R}^n} (\nabla f(x))\,(\nabla f(x))^t\,dx,$$

where $\nabla f(x)$ is written as a column vector and $(\nabla f(x))^t$ denotes its transpose, and hence it follows from (4) that for $s \in \mathbb{R}$ and $\Phi(x) = \phi x + y$, where $\phi \in \mathrm{GL}(n)$ and $y \in \mathbb{R}^n$,

$$\begin{aligned} \mathrm{J}((s\,f)^2 \circ \Phi^{-1}) &= 4\,s^2 \int_{\mathbb{R}^n} (\phi^{-t}\,\nabla f(\Phi^{-1}))\,(\phi^{-t}\,\nabla f(\Phi^{-1}))^t\,dx \\ &= s^2\,|\det \phi|\;\phi^{-t}\,\mathrm{J}(f^2)\,\phi^{-1}. \end{aligned}$$

Thus we have

$$\mathrm{J}((s\,f)^2) = s^2\,\mathrm{J}(f^2),\;\; \mathrm{J}(f^2 \circ \phi^{-1}) = |\det \phi|\;\phi^{-t}\,\mathrm{J}(f^2)\phi^{-1}$$

and

$$\mathrm{J}(f^2 \circ \tau^{-1}) = \mathrm{J}(f)$$

for all $s \in \mathbb{R}$, translations τ and $\phi \in \mathrm{GL}(n)$. Consequently, the operator $f \mapsto c\,\mathrm{J}(f^2)$ is affinely contravariant. $\square$

The following lemma establishes an important connection between matrix-valued operators on $W^{1,2}(\mathbb{R}^n)$ and on $\mathcal{P}_0^n$. For a further such connection, see [18, Lemma 5].

Lemma 3. *For $P \in \mathcal{P}_0^n$, we have $\mathrm{J}(\ell_P^2) = \frac{4}{n}\,\mathrm{L}(P)$.*

Proof. Let $\mathcal{N}(P)$ denote the finite set of unit outer normal vectors to facets of P. Let $F(u)$ be the facet with normal vector $u \in \mathcal{N}(P)$. Let $a(P,u)$ the $(n-1)$-dimensional area of $F(u)$, $h(P,u)$ the distance from the hyperplane containing $F(u)$ to the origin, and $T(u)$ the convex hull of $F(u)$ and the origin. Since for $x \in T(u)$

$$\ell_P(x) = -\frac{u}{h(P,u)} \cdot x + 1$$

and

$$\frac{\partial \ell_P}{\partial x_i}(x) = -\frac{u_i}{h(P,u)},$$

we obtain that

$$
\begin{aligned}
\mathrm{J}_{ij}(\ell_P^2) &= 4\int_{\mathbb{R}^n} \frac{\partial \ell_P(x)}{\partial x_i}\,\frac{\partial \ell_P(x)}{\partial x_j}\,dx \\
&= 4\sum_{u\in\mathcal{N}(P)} \frac{u_i\,u_j}{h(P,u)^2}\int_{T(u)} dx \\
&= \frac{4}{n}\sum_{u\in\mathcal{N}(P)} \frac{a(P,u)}{h(P,u)}\,u_i\,u_j.
\end{aligned}
$$

Combined with definition (3), this concludes the proof of the lemma. □

4 Proof of the Theorem

In Lemma 2, it was shown that $f\mapsto c\,\mathrm{J}(f^2)$ is a continuous and affinely contravariant valuation on $W^{1,2}(\mathbb{R}^n)$. Now suppose that $\mathrm{Z}: W^{1,2}(\mathbb{R}^n)\to\langle\mathbb{M}^n,+\rangle$ is a continuous and affinely contravariant valuation. The proof that there is a constant $c\in\mathbb{R}$ such that $\mathrm{Z}(f)=c\,\mathrm{J}(f^2)$ for all $f\in W^{1,2}(\mathbb{R}^n)$ consists of several steps. First, we show that the weight of Z is greater or equal to 1. In Lemma 5, we combine this with Theorem 1 and Lemma 3 and show that there is a constant $c\in\mathbb{R}$ such that $\mathrm{Z}(f)=c\,\mathrm{J}(f^2)$ for all $f\in P^{1,2}(\mathbb{R}^n)$. In particular, Z is $\mathrm{GL}(n)$ contravariant of weight $p=1$. In Lemma 6, we show that this implies that if Z is non-trivial on $P^{1,2}(\mathbb{R}^n)$, then it is homogeneous of degree 2. Here Z is called trivial on $P^{1,2}(\mathbb{R}^n)$ if $\mathrm{Z}(f)=0$ for all $f\in P^{1,2}(\mathbb{R}^n)$. In the last step, we show that every homogeneous, continuous, and translation invariant valuation is already determined by its values on $P^{1,2}(\mathbb{R}^n)$. Combined with Lemmas 5 and 6 this completes the proof of the theorem.

Lemma 4. *If* $\mathrm{Z}: L^{1,2}(\mathbb{R}^n)\to\mathbb{M}^n$ *is continuous, non-trivial on* $P^{1,2}(\mathbb{R}^n)$ *and* $\mathrm{GL}(n)$ *contravariant of weight* p, *then* $p\geq 1$.

Proof. For $a>0$ and $0<\varepsilon<1$, define $\phi_a\in\mathrm{GL}(n)$ by $\phi_a e_i = a\,e_i$ for $i=1,2$ and by $\phi_a e_i = a^{\varepsilon} e_i$ for $i=3,\dots,n$. For $P\in\mathcal{P}_0^n$, we have by (8)

$$
\|\ell^2_{\phi_a P}\|_2^2 = \int_{\mathbb{R}^n} \ell_P^4(\phi_a^{-1}x)\,dx = |\det\phi_a|\int_{\mathbb{R}^n}\ell_P^4(x)\,dx = O(a^{2+(n-2)\varepsilon})
$$

and by (8) and (4)

$$\begin{aligned}
\|\nabla \ell^2_{\phi_a P}\|_2^2 &= \int_{\mathbb{R}^n} |\nabla(\ell_P^2 \circ \phi_a^{-1})(x)|^2\,dx \\
&= \int_{\mathbb{R}^n} |\phi^{-t}\,\nabla \ell_P^2(\phi_a^{-1}x)|^2\,dx \\
&= |\det \phi_a| \int_{\mathbb{R}^n} |\phi^{-t}\,\nabla \ell_P^2(x)|^2\,dx \\
&= O(a^{2+(n-2)\varepsilon} \max_{u\in S^{n-1}} |\phi_a^{-t}u|^2) \\
&= O(a^{(n-2)\varepsilon})
\end{aligned}$$

as $a \to 0$. Hence $\ell^2_{\phi_a P} \to 0$ in $W^{1,2}(\mathbb{R}^n)$ as $a \to 0$. Since Z is $\mathrm{GL}(n)$ contravariant of weight p,

$$\mathrm{Z}(\ell^2_{\phi_a P}) = a^{(2+(n-1)\varepsilon)p} \phi_a^{-t}\,\mathrm{Z}(\ell_P^2)\phi_a^{-1}.$$

Thus

$$\mathrm{Z}_{11}(\ell^2_{\phi_a P}) = a^{(2+(n-1)\varepsilon)p-2}\,\mathrm{Z}_{11}(\ell_P^2). \tag{10}$$

Since Z is non-trivial on $P^{1,2}(\mathbb{R}^n)$, there is a polytope $Q \in \mathcal{P}_0^n$ such that $\mathrm{Z}(\ell_Q^2) \neq 0$. Since $\mathrm{Z}(\ell_Q^2)$ is symmetric, there is an orthogonal transformation ψ such that $\psi^{-t}\,\mathrm{Z}(\ell_Q^2)\psi^{-1}$ is a diagonal matrix. Since Z is $\mathrm{GL}(n)$ contravariant, we see that $\mathrm{Z}(\ell^2_{\psi Q})$ is a diagonal matrix. This shows that after exchanging the coordinates if necessary we can choose $P \in \mathcal{P}_0^n$ such that $\mathrm{Z}_{11}(\ell_P^2) \neq 0$. Since Z is continuous, (10) and (5) imply that $p \geq 2/(2+(n-1)\varepsilon)$. Since $\varepsilon > 0$ was arbitrary, we obtain $p \geq 1$. $\square$

Lemma 5. *If* $\mathrm{Z} : P^{1,2}(\mathbb{R}^n) \to \langle \mathbb{M}^n, + \rangle$ *is a continuous and affinely contravariant valuation, then there is a constant* $c \geq 0$ *such that*

$$\mathrm{Z}(f) = c\,\mathrm{J}(f^2)$$

for every $f \in P^{1,2}(\mathbb{R}^n)$.

Proof. Define the operator $\mathrm{Y} : \mathcal{P}_0^n \to \langle \mathbb{M}^n, + \rangle$ by setting

$$\mathrm{Y}(P) = \mathrm{Z}(\ell_P^2).$$

If $\ell_P, \ell_Q \in P^{1,2}(\mathbb{R}^n)$ are such that $\ell_P \vee \ell_Q \in P^{1,2}(\mathbb{R}^n)$, then $\ell_P \vee \ell_Q = \ell_{P\cup Q}$ and $\ell_P \wedge \ell_Q = \ell_{P\cap Q}$. Since Z is a valuation on $P^{1,2}(\mathbb{R}^n)$, it follows that for

$P, Q, P \cup Q \in \mathcal{P}_0^n$,

$$\begin{aligned} \mathrm{Y}(P) + \mathrm{Y}(Q) &= \mathrm{Z}(\ell_P^2) + \mathrm{Z}(\ell_Q^2) \\ &= \mathrm{Z}(\ell_P^2 \vee \ell_Q^2) + \mathrm{Z}(\ell_P^2 \wedge \ell_Q^2) \\ &= \mathrm{Y}(P \cup Q) + \mathrm{Y}(P \cap Q). \end{aligned}$$

Thus $\mathrm{Y} : \mathcal{P}_0^n \to \langle \mathbb{M}^n, + \rangle$ is a valuation.

By Lemma 4, Z is $\mathrm{GL}(n)$ contravariant of weight $p \geq 1$. Since for $\phi \in \mathrm{GL}(n)$ we have by (8)

$$\mathrm{Y}(\phi P) = \mathrm{Z}(\ell_P^2 \circ \phi^{-1}) = |\det \phi|^p \ \phi^{-t}\, \mathrm{Z}(\ell_P^2) \phi^{-1} = |\det \phi|^p \ \phi^{-t} (\mathrm{Y}\, P) \phi^{-1},$$

also Y is $\mathrm{GL}(n)$ contravariant of weight $p \geq 1$. Thus by Theorem 1 there exists a constant $c \in \mathbb{R}$ such that

$$\mathrm{Z}(\ell_P^2) = c\, \mathrm{L}(P)$$

for all $\ell_P \in P^{1,2}(\mathbb{R}^n)$. The statement now follows from Lemma 3. □

Lemma 6. *If* $\mathrm{Z} : W^{1,2}(\mathbb{R}^n) \to \langle \mathbb{M}^n, + \rangle$ *is a continuous and non-trivial valuation which is affinely contravariant of weight 1 and homogeneous of degree* q*, then* $q = 2$*.*

Proof. First, we show that $q \geq 2$. Let $P \in \mathcal{P}_0^n$ and $\varepsilon > 0$. Take translations $\tau_1, \dots, \tau_{k^2}$ such that the polytopes $\tau_i\, P$ are pairwise disjoint. Define

$$f_k = \frac{1}{k^{1+\varepsilon}} (\ell_{\tau_1 P}^2 \vee \cdots \vee \ell_{\tau_{k^2} P}^2).$$

Then $\|f_k\|_2 = \|\nabla f_k\|_2 = O(k^{-\varepsilon})$ as $k \to \infty$. Hence $f_k \to 0$ as $k \to \infty$ in $W^{1,2}(\mathbb{R}^n)$. Using (5), the valuation property and translation invariance of Z, we see that

$$\mathrm{Z}(\ell_{\tau_1 P}^2 \vee \ell_{\tau_2 P}^2) = \mathrm{Z}(\ell_{\tau_1 P}^2 \vee \ell_{\tau_2 P}^2) + \mathrm{Z}(\ell_{\tau_1 P}^2 \wedge \ell_{\tau_2 P}^2) = \mathrm{Z}(\ell_{\tau_1 P}^2) + \mathrm{Z}(\ell_{\tau_2 P}^2) = 2\, \mathrm{Z}(\ell_P^2).$$

Hence

$$\mathrm{Z}(f_k) = k^2\, k^{-q(1+\varepsilon)}\, \mathrm{Z}(\ell_P^2).$$

Since Z is continuous and $f_k \to 0$ in $W^{1,2}(\mathbb{R}^n)$, combined with (5) this implies that $q \geq 2$.

Next, we show that $q \leq 2$. Let $P \in \mathcal{P}_0^n$ and $\alpha, \beta > 0$. Take translations $\tau_1, \dots, \tau_k$ such that the polytopes $\tau_i\, P$ are pairwise disjoint. Define

$$f_k = k^\alpha (\ell_{\tau_1 (P/k^\beta)}^2 \vee \cdots \vee \ell_{\tau_k (P/k^\beta)}^2).$$

We write ϕ for $n\times n$ identity matrix multiplied with $k^{-\beta}$ and have $\ell^2_{P/k^\beta} = \ell^2_{\phi P} = \ell^2_P\circ\phi^{-1}$. Hence

$$\|\ell^2_{\phi P}\|_2^2 = \int_{\phi P} \ell^4_P(\phi^{-1}x)\,dx = k^{-n\beta}\int_P \ell^4_p(x)\,dx$$

and by (4),

$$\begin{aligned}\|\nabla\ell^2_{\phi P}\|_2^2 &= \int_{\phi P} |\nabla(\ell^2_P\circ\phi^{-1})(x)|^2\,dx\\ &= \int_{\phi P} |\phi^{-t}\nabla\ell^2_P(\phi^{-1}x)|^2\,dx\\ &= k^{-(n-2)\beta}\int_P |\nabla\ell^2_P(x)|^2\,dx.\end{aligned}$$

Hence $\|f_k\|_2 = O(k^{1+\alpha-n\,\beta/2})$ and $\|\nabla f_k\|_2 = O(k^{1+\alpha-(n-2)\,\beta/2})$ as $k\to\infty$. Let $\alpha<(n-2)\beta/2-1$. Then we obtain that $f_k\to 0$ as $k\to\infty$ in $W^{1,2}(\mathbb{R}^n)$. Since Z is $\mathrm{GL}(n)$ contravariant of weight 1, we have

$$\mathrm{Z}(\ell^2_{P/k^\beta}) = \mathrm{Z}(\ell^2_p\circ\phi^{-1}) = k^{-(n-2)\beta}\,\mathrm{Z}(\ell^2_P).$$

Hence

$$\mathrm{Z}(f_k) = k\,k^{\alpha\,q}k^{-(n-2)\beta}\,\mathrm{Z}(\ell^2_P).$$

Since Z is continuous and $f_k\to 0$ in $W^{1,2}(\mathbb{R}^n)$, it follows from (5) that $q<(-1+(n-2)\beta)/\alpha$. Since this holds for all $\alpha<(n-2)\beta/2-1$, letting $\beta\to\infty$ gives $q\le 2$. $\square$

The following lemma on valuations on $L^{1,2}(\mathbb{R}^n)$, the space of piecewise affine functions, is proved similarly to Lemma 9 in [32].

Lemma 7. *Let $\mathrm{Z}_1,\mathrm{Z}_2 : L^{1,2}(\mathbb{R}^n)\to\langle\mathbb{M}^n,+\rangle$ be continuous and translation invariant valuations, which are homogeneous of the same degree. If $\mathrm{Z}_1(f) = \mathrm{Z}_2(f)$ holds for all $f\in P^{1,2}(\mathbb{R}^n)$, then*

$$\mathrm{Z}_1(f) = \mathrm{Z}_2(f) \tag{11}$$

for all $f\in L^{1,2}(\mathbb{R}^n)$.

Proof. Let Z_1 and Z_2 be homogeneous of degree q. As noted in Section 2, it is clear that $q\ge 0$ and that if $q=0$, then Z_1 and Z_2 are constant. Hence (11) holds for $q=0$ and we assume that $q>0$. Thus we know that

$\mathrm{Z}_1(0) = \mathrm{Z}_2(0) = 0$. Since Z_1 and Z_2 are valuations and homogeneous, this implies that for $i = 1, 2$,

$$\mathrm{Z}_i(f \vee 0) + \mathrm{Z}_i(f \wedge 0) = \mathrm{Z}_i(f) + \mathrm{Z}_i(0) = \mathrm{Z}_i(f)$$

and

$$\mathrm{Z}_i(f \wedge 0) = \mathrm{Z}_i(-((-f) \vee 0)) = \mathrm{Z}_i((-f) \vee 0).$$

Thus it suffices to show that (11) holds for all $f \in L^{1,2}(\mathbb{R}^n)$ with $f \geq 0$.

Let such a function f be given and let f not vanish identically. Triangulate the support of f so that f is affine on each simplex of the triangulation. Let V be the (finite) set of vertices and $\mathcal{S}$ the set of n-dimensional simplices of this triangulation. Note that f is determined by the values $f(v)$ for $v \in V$ and that if $f(\bar{v}) > 0$ for some $\bar{v} \in V$, then by changing the value $f(\bar{v})$ we obtain again a function in $L^{1,2}(\mathbb{R}^n)$. Since Z_1 and Z_2 are continuous, it suffices to prove (11) for a function f where the values $f(v)$ are distinct for $v \in V$ with $f(v) > 0$.

First, we show that there are functions $f_1, \dots, f_m \in L^{1,2}(\mathbb{R}^n)$ which are non-negative and concave on their supports such that

$$f = f_1 \vee \dots \vee f_m. \tag{12}$$

Let S_i be a simplex of $\mathcal{S}$ and define the function f_i by setting $f_i(v) = f(v)$ on the vertices v of the simplex S_i. Choose a convex polytope P_i such that $S_i \subset P_i$ and set $f_i(v) = 0$ on the vertices v of P_i. The function f_i determined by this data is concave on its support and piecewise linear. Moreover, if the polytopes P_i are chosen suitably small, (12) holds.

Using the inclusion-exclusion principle and (12), we obtain that for $i = 1, 2$,

$$\mathrm{Z}_i(f) = \mathrm{Z}_i(f_1 \vee \dots \vee f_m) = \sum\nolimits_J (-1)^{|J|-1}\, \mathrm{Z}_i(f_J)$$

where we sum over all non-empty $J \subset \{1, \dots, m\}$,

$$f_J = f_{j_1} \wedge \dots \wedge f_{j_k}$$

for $J = \{j_1, \dots, j_k\}$, and $|J|$ is the cardinality of J. Note that the functions f_J are concave on their support. Thus it suffices to prove (11) for $f \in L^{1,2}(\mathbb{R}^n)$ such that $f \geq 0$ and f is concave on its support.

For a given function $f \in L^{1,2}(\mathbb{R}^n)$ which is concave on its support, let $F \subset \mathbb{R}^{n+1}$ be the compact polytope bounded by the graph of f and the hyperplane $\{x_{n+1} = 0\}$. We call F *singular* if F has n facet hyperplanes that intersect in a line L parallel to $\{x_{n+1} = 0\}$ but not contained in $\{x_{n+1} = 0\}$.

Since Z_1 and Z_2 are continuous, it suffices to show (11) for $f \in L^{1,2}(\mathbb{R}^n)$ such that F is not singular. So we assume for the rest of the proof that f has this property.

Let such a function f be given. Let $\bar{p}$ be the vertex of F with the largest x_{n+1} coordinate. We use induction on the number m of facet hyperplanes of F that are not passing through $\bar{p}$. If $m = 1$, then a translate of f is in $P^{1,2}(\mathbb{R}^n)$. Since Z_1 and Z_2 are translation invariant, (11) is true. Suppose (11) is true for all $f \in L^{1,2}(\mathbb{R}^n)$ with at most $(m-1)$ facet hyperplanes not containing $\bar{p}$. We show that (11) then also holds for all $f \in L^{1,2}(\mathbb{R}^n)$ with m such hyperplanes.

So let F have m such hyperplanes. Let $p_0 = (x_0, f(x_0))$ be the vertex of F with minimal non-negative x_{n+1}-coordinate. Let $H_1, \dots, H_k$ be the facet hyperplanes of F through p_0 which do not contain $\bar{p}$. There is at least one such hyperplane. Define $\bar{F}$ as the polytope bounded by all facet hyperplanes of F with the exception of $H_1, \dots, H_k$. Since F has no edges parallel to $\{x_{n+1} = 0\}$ but not contained in $\{x_{n+1} = 0\}$, $\bar{F}$ is bounded and the function $\bar{f}$ corresponding to $\bar{F}$ is in $L^{1,2}(\mathbb{R}^n)$. Note that the graph of $\bar{f}$ has at most $(m-1)$ facet hyperplanes not containing $\bar{p}$. Let $\bar{H}_1, \dots, \bar{H}_i$ be the facet hyperplanes of $\bar{F}$ that contain p_0. Choose suitable hyperplanes $\bar{H}_{i+1}, \dots, \bar{H}_k$ containing p_0 so that the hyperplanes $\bar{H}_1, \dots, \bar{H}_k$ and $\{x_{n+1} = 0\}$ bound a pyramid with apex at p_0 that is contained in $\bar{F}$, has x_0 in its base and has $\bar{H}_1, \dots, \bar{H}_i$ among its facet hyperplanes. Define $\bar{\ell}$ as the piecewise affine function determined by this pyramid and note that a suitable translate of $\bar{\ell}$ is in $P^{1,2}(\mathbb{R}^n)$. Let $\ell = f \wedge \bar{\ell} \in L^{1,2}(\mathbb{R}^n)$. The polytope determined by ℓ is a pyramid since it is bounded by $\{x_{n+1} = 0\}$ and hyperplanes containing p_0. Therefore a suitable translate of ℓ is in $P^{1,2}(\mathbb{R}^n)$. Since Z_i is a valuation and

$$f \vee \bar{\ell} = \bar{f} \quad \text{and} \quad f \wedge \bar{\ell} = \ell,$$

we have

$$\mathrm{Z}_i(f) + \mathrm{Z}_i(\bar{\ell}) = \mathrm{Z}_i(f \vee \bar{\ell}) + \mathrm{Z}_i(f \wedge \bar{\ell}) = \mathrm{Z}_i(\bar{f}) + \mathrm{Z}_i(\ell). \tag{13}$$

Since translates of ℓ and $\bar{\ell}$ are in $P^{1,2}(\mathbb{R}^n)$, by assumption $\mathrm{Z}_1(\bar{\ell}) = \mathrm{Z}_2(\bar{\ell})$ and $\mathrm{Z}_1(\ell) = \mathrm{Z}_2(\ell)$. Since the polytope $\bar{F}$ has at most $(m-1)$ facet hyperplanes not containing $\bar{p}$, by induction $\mathrm{Z}_1(\bar{f}) = \mathrm{Z}_2(\bar{f})$. Thus (13) implies that (11) holds for all $f \in L^{1,2}(\mathbb{R}^n)$ with m facet hyperplanes not containing $\bar{p}$. This completes the proof of the lemma. $\square$

References

[1] S. Alesker, *Continuous rotation invariant valuations on convex sets*, Ann. of Math. **149** (1999), 977–1005.

[2] S. Alesker, *Description of translation invariant valuations on convex sets with solution of P. McMullen's conjecture*, Geom. Funct. Anal. **11** (2001), 244–272.

[3] S. Alesker, *Hard Lefschetz theorem for valuations, complex integral geometry, and unitarily invariant valuations*, J. Differential Geom. **63** (2003), 63–95.

[4] S. Alesker, *The multiplicative structure on continuous polynomial valuations*, Geom. Funct. Anal. **14** (2004), 1–26.

[5] S. Alesker, *Valuations on convex sets, non-commutative determinants, and pluripotential theory*, Adv. Math. **195** (2005), 561–595.

[6] S. Alesker, *Theory of valuations of manifolds, II*, Adv. Math. **207** (2006), 420–454.

[7] S. Alesker, *Theory of valuations on manifolds: a survey*, Geom. Funct. Anal. **17** (2007), 1321–1341.

[8] S. Alesker and J. H. G. Fu, *Theory of valuations on manifolds. III. Multiplicative structure in the general case*, Trans. Amer. Math. Soc. **360** (2008), 1951–1981.

[9] A. Bernig, *Valuations with Crofton formula and Finsler geometry*, Adv. Math. **210** (2007), 733–753.

[10] A. Bernig and L. Bröcker, *Valuations on manifolds and Rumin cohomology*, J. Differential Geom. **75** (2007), 433–457.

[11] A. Bernig and J.H.G. Fu, *Hermitian integral geometry*, Ann. of Math., in press.

[12] S. Campi and P. Gronchi, *The L^p-Busemann-Petty centroid inequality*, Adv. Math. **167** (2002), 128–141.

[13] K. Chou and X.-J. Wang, *The L_p-Minkowski problem and the Minkowski problem in centroaffine geometry*, Adv. Math. **205** (2006), 33–83.

[14] T. Cover and J. Thomas, *Elements of information theory*, second ed., Wiley-Interscience, Hoboken, NJ, 2006.

[15] A. Dembo, T. Cover, and J. Thomas, *Information-theoretic inequalities*, IEEE Trans. Inform. Theory **37** (1991), 1501–1518.

[16] J.H.G. Fu, *Structure of the unitary valuation algebra*, J. Differential Geom. **72** (2006), 509–533.

[17] E. Grinberg and G. Zhang, *Convolutions, transforms, and convex bodies*, Proc. London Math. Soc. **78** (1999), 77–115.

[18] O. Guleryuz, E. Lutwak, D. Yang, and G. Zhang, *Information-theoretic inequalities for contoured probability distributions*, IEEE Trans. Inform. Theory **48** (2002), 2377–2383.

[19] C. Haberl, *Star body valued valuations*, Indiana Univ. Math. J. **58** (2009), 2253–2276.

[20] C. Haberl, *Blaschke valuations*, Amer. J. Math., in press.

[21] C. Haberl, *Minkowski valuations intertwining the special linear group*, preprint.

[22] C. Haberl and M. Ludwig, *A characterization of L_p intersection bodies*, Int. Math. Res. Not. **10548** (2006), 1–29.

[23] C. Haberl and F. Schuster, *General L_p affine isoperimetric inequalities*, J. Differential Geom. **83** (2009), 1–26.

[24] H. Hadwiger, *Vorlesungen über Inhalt, Oberfläche und Isoperimetrie*, Springer, Berlin, 1957.

[25] D. A. Klain, *Star valuations and dual mixed volumes*, Adv. Math. **121** (1996), 80–101.

[26] D. A. Klain, *Invariant valuations on star-shaped sets*, Adv. Math. **125** (1997), 95–113.

[27] D. A. Klain and G.-C. Rota, *Introduction to geometric probability*, Cambridge University Press, Cambridge, 1997.

[28] M. Ludwig, *Projection bodies and valuations*, Adv. Math. **172** (2002), 158–168.

[29] M. Ludwig, *Ellipsoids and matrix valued valuations*, Duke Math. J. **119** (2003), 159–188.

[30] M. Ludwig, *Intersection bodies and valuations*, Amer. J. Math. **128** (2006), 1409–1428.

[31] M. Ludwig, *Minkowski areas and valuations*, J. Differential Geom., in press.

[32] M. Ludwig, *Valuations on Sobolev spaces*, preprint.

[33] M. Ludwig and M. Reitzner, *A characterization of affine surface area*, Adv. Math. **147** (1999), 138–172.

[34] M. Ludwig and M. Reitzner, *A classification of $SL(n)$ invariant valuations*, Ann. of Math., in press.

[35] E. Lutwak, *The Brunn-Minkowski-Firey theory. I. Mixed volumes and the Minkowski problem*, J. Differential Geom. **38** (1993), 131–150.

[36] E. Lutwak, *The Brunn-Minkowski-Firey theory. II: Affine and geominimal surface areas*, Adv. Math. **118** (1996), 244–294.

[37] E. Lutwak, D. Yang, and G. Zhang, *A new ellipsoid associated with convex bodies*, Duke Math. J. **104** (2000), 375–390.

[38] E. Lutwak, D. Yang, and G. Zhang, *L_p affine isoperimetric inequalities*, J. Differential Geom. **56** (2000), 111–132.

[39] E. Lutwak, D. Yang, and G. Zhang, *The Cramer-Rao inequality for star bodies*, Duke Math. J. **112** (2002), 59–81.

[40] E. Lutwak, D. Yang, and G. Zhang, *Moment-entropy inequalities*, Ann. Probab. **32** (2004), 757–774.

[41] E. Lutwak, D. Yang, and G. Zhang, *Volume inequalities for subspaces of L_p*, J. Differential Geom. **68** (2004), 159–184.

[42] E. Lutwak and G. Zhang, *Blaschke-Santaló inequalities*, J. Differential Geom. **47** (1997), 1–16.

[43] R. Schneider and F. Schuster, *Rotation equivariant Minkowski valuations*, Int. Math. Res. Not. **72894** (2006), 1–20.

[44] F. Schuster, *Valuations and Busemann-Petty type problems*, Adv. Math. **219** (2008), 344–368.

[45] F. Schuster, *Crofton measures and Minkowski valuations*, Duke Math. J. **154** (2010), 1–30.

[46] F. Schuster and T. Wannerer, *GL(n) contravariant Minkowski valuations*, Trans. Amer. Math. Soc., in press.

[47] A. Stancu, *The discrete planar L_0-Minkowski problem*, Adv. Math. **167** (2002), 160–174.

[48] A. Stancu, *On the number of solutions to the discrete two-dimensional L_0-Minkowski problem*, Adv. Math. **180** (2003), 290–323.

Institut für Diskrete Mathematik und Geometrie, TU Wien,
Wiedner Hauptstr. 8-10/104, 1040 Wien, Austria

and

Department of Mathematics, Polytechnic Institute of New York University,
6 MetroTech Center, Brooklyn, NY 11201, U.S.A.

E-mail: monika.ludwig@tuwien.ac.at