Orthogonality in Generalized Minkowski Spaces

Thomas Jahn
Faculty of Mathematics, Technische Universität Chemnitz
09107 Chemnitz, Germany
thomas.jahn@mathematik.tu-chemnitz.de

We combine functional analytic and geometric viewpoints on approximate Birkhoff and isosceles orthogonality in generalized Minkowski spaces which are finite-dimensional vector spaces equipped with a gauge. Thus we present the first approach to orthogonality types in such spaces.

Keywords: best approximation, bisector, Birkhoff orthogonality, gauge, generalized Minkowski space, isosceles orthogonality
MSC(2010): 46B20, 52A20, 52A21, 52A41, 90C25

1 Introduction

In recent literature, some attention was attracted to the investigation of geometric properties of generalized Minkowski spaces defined by so-called gauges. In finite dimensions, these are positively homogeneous and subadditive functions $\gamma : \mathbb{R}^d \to \mathbb{R}$ which have non-negative values and vanish only at the origin $0 \in \mathbb{R}^d$. Obviously, this notion is a generalization of that of a norm because only the homogeneity property has been relaxed. The good news is that many of the concepts of classical functional analysis of finite-dimensional normed spaces still work in generalized Minkowski spaces $(\mathbb{R}^d, \gamma)$, see [16].

A convenient way for understanding such spaces is to exploit the correspondence between analysis of gauges and the geometry of their unit balls. Examples for this interplay can be found in [32] and [33]. The purpose of the present paper is to add another subject to this list, namely that of orthogonality.

Motivated by the pleasant theory of Hilbert spaces, mathematicians introduced various generalizations of the notion of orthogonality for non-Hilbert spaces. Birkhoff orthogonality is the most popular one, and its usage in the setting of normed spaces reaches from angular measures [14], approximation theory [29], curve theory [44, 58], orthocentric systems [49], matrix theory [8, 9], and orthogonal decompositions of Banach spaces [1, 11, 37] to random processes [53, § 2.9]. Although it is named after Birkhoff [10], there are earlier
papers by Radon [52] and Blaschke [12] on this orthogonality type. Properties of isosceles orthogonality, a notion introduced by James [34], are determined by the geometry of bisectors. These sets of points of equal distance from given objects have a great impact on Computational Geometry and Discrete Geometry by being a main tool in creating Voronoi diagrams; see, for instance, [19] for recent research in this direction.

In order to take the next step and transfer these concepts to generalized Minkowski spaces, we introduce binary relations \( \perp \) on \( \mathbb{R}^d \) via proximality with respect to certain convex functions. Therefore, we fix our notation and outline necessary concepts from convex analysis and convex optimization in the second section. In the classical theory of normed spaces, the presence of the following properties indicates “how much” a given orthogonality notion resembles usual Euclidean orthogonality:

(a) Non-degeneracy: For all \( x \in \mathbb{R}^d \) and \( \lambda, \mu \in \mathbb{R} \), \( \lambda x \perp \mu x \) if and only if \( \lambda \mu x = 0 \).
(b) Symmetry: For all \( x, y \in \mathbb{R}^d \), \( x \perp y \) implies \( y \perp x \).
(c) Right additivity: For all \( x, y, z \in \mathbb{R}^d \), \( x \perp y \) and \( x \perp z \) together imply \( x \perp (y + z) \).
(d) Left additivity: For all \( x, y, z \in \mathbb{R}^d \), \( x \perp z \) and \( y \perp z \) together imply \( (x + y) \perp z \).
(e) Right homogeneity: For all \( x, y \in \mathbb{R}^d \) and \( \lambda > 0 \), \( x \perp y \) implies \( \lambda x \perp \lambda y \).
(f) Left homogeneity: For all \( x, y \in \mathbb{R}^d \) and \( \lambda > 0 \), \( x \perp y \) implies \( \lambda x \perp y \).
(g) Right existence: For all \( x, y \in \mathbb{R}^d \), there exists a number \( \alpha \in \mathbb{R} \) such that \( x \perp (\alpha x + y) \).
(h) Left existence: For all \( x, y \in \mathbb{R}^d \), there exists a number \( \alpha \in \mathbb{R} \) such that \( (\alpha x + y) \perp x \).

Checking these properties for approximate Birkhoff orthogonality and isosceles orthogonality in generalized Minkowski spaces is (sometimes more, sometimes less explicitly) the subject of the sections 3 and 4. We conclude the presentation by giving some future research perspectives in the fifth and final section.

2 Notation and preliminaries

Throughout the paper, we are concerned with the \( d \)-dimensional real vector space \( \mathbb{R}^d \). This space is equipped with the usual topology induced by the Euclidean inner product \( \langle \cdot, \cdot \rangle \) and the Euclidean norm \( \| \cdot \| \). A gauge is a function \( \gamma : \mathbb{R}^d \to \mathbb{R} \) which meets the following requirements:

(a) \( \gamma(x) \geq 0 \) for all \( x \in \mathbb{R}^d \), and \( \gamma(x) = 0 \) implies \( x = 0 \),
(b) \( \gamma(\lambda x) = \lambda \gamma(x) \) for all \( x \in \mathbb{R}^d \), \( \lambda > 0 \),
(c) \( \gamma(x + y) \leq \gamma(x) + \gamma(y) \).

In particular, \( \gamma \) is called rotund if \( \gamma(x + y) < 2 \) whenever \( x, y \in \mathbb{R}^d \) and \( \gamma(x) = \gamma(y) = 1 \). It is called Gâteaux differentiable at \( x \in \mathbb{R}^d \), if the directional derivative

\[
\gamma'(x; y) = \lim_{\lambda \to 0} \frac{\gamma(x + \lambda y) - \gamma(x)}{\lambda}
\]  

is linear in \( y \). The polar function of \( \gamma \) is defined as

\[
\gamma^\circ : \mathbb{R}^d \to \mathbb{R}, \quad \gamma^\circ(x^*) = \inf \left\{ \lambda > 0 \mid \langle x^* , x \rangle \leq \lambda \gamma(x) \forall x \in \mathbb{R}^d \right\}.
\]
The polar function $\gamma^*$ is again a gauge and satisfies the Cauchy–Schwarz inequality
\[ \langle x^* | x \rangle \leq \gamma^*(x^*)\gamma(x) \tag{2} \]
automatically.

A set $B \subseteq \mathbb{R}^d$ is convex if the line segment $[x, y] := (\lambda x + (1 - \lambda)y | 0 \leq \lambda \leq 1)$ is contained in $B$ for all $x, y \in B$. A convex set $B$ is rotund if there is no line segment contained in its boundary. For $x^* \in \mathbb{R}^d$ and $a \in \mathbb{R}$, the hyperplane $\{y \in \mathbb{R}^d | \langle x^* | y \rangle = a\}$ is a supporting hyperplane of $B$ provided $B$ is contained in the closed half-space $\{y \in \mathbb{R}^d | \langle x^* | y \rangle \leq a\}$ but not in its interior. The set $B$ is called smooth if, for every boundary point $x$ of $B$, there is a unique supporting hyperplane of $B$ passing through $x$. An affine diameter of $B$ is a line segment $[x, y]$ joining two boundary points of $B$ which admit parallel supporting hyperplanes passing through them. The support function of $B$ is $h_B(x^*) = \sup \{\langle x^* | x \rangle | x \in \mathbb{R}^d, \gamma(x) \leq 1\}$, and the set $\{x^* \in \mathbb{R}^d | h_B(x^*) \leq 1\}$ is the polar set of $B$. A set $K \subseteq \mathbb{R}^d$ is a cone provided $\lambda K = K$ for all $\lambda > 0$. The formulas $\gamma(x) = \inf(\lambda > 0 | x \in \lambda B)$ and $B = \{x \in \mathbb{R}^d | \gamma(x) \leq 1\}$ establish a one-to-one correspondence between gauges and compact and convex sets having the origin as interior point. Analogously to the classical theory of normed spaces, we define the ball with radius $\lambda$ and center $x$ by $B(x, \lambda) = \{y \in \mathbb{R}^d | \gamma(y - x) \leq \lambda\}$. In particular, $B(0, 1)$ is the unit ball of the generalized Minkowski space $(\mathbb{R}^d, \gamma)$. The intimate relationship between gauges and unit balls goes further: We will see (Theorems 3.12 and 3.11) that rotund gauges have rotund unit balls, Gâteaux differentiable gauges have smooth unit balls, and vice versa. Polarity also behaves well in the interplay between gauges and unit balls. Namely, the polar gauge $\gamma^*$ coincides with the support function $h_{B(0,1)}$, and its unit ball is the polar set of $B(0,1)$. Therefore, the polar gauge of a norm is again a norm (known as the dual norm).

It is known that convex functions $f : \mathbb{R}^d \rightarrow \mathbb{R}$—that is $f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y)$ for all $x, y \in \mathbb{R}^d$ and $\lambda \in \mathbb{R}$—are continuous and their $\varepsilon$-subdifferential
\[
\partial_\varepsilon f(x) := \left\{x^* \in \mathbb{R}^d \left| \langle x^* | y - x \rangle \leq f(y) - f(x) + \varepsilon \text{ for all } y \in \mathbb{R}^d \right. \right\}
\]
is non-empty, compact, and convex at each point $x \in \mathbb{R}^d$, see [6, Corollary 8.30] and [63, Theorem 2.4.2(ii)]. The $\varepsilon$-directional derivative
\[
f'_\varepsilon(x; y) := \inf_{\lambda > 0} \frac{f(x + \lambda y) - f(x) + \varepsilon}{\lambda}
\]
coincides with the support function of $\partial_\varepsilon(x)$, see [63, Theorem 2.1.14, Theorem 2.4.9]. We shall omit $\varepsilon$ from the notation if it equals zero, that is $f' := f'_0$ and $\partial f := \partial_0 f$. Note that every subgradient $x^* \in \partial f(x)$ defines a supporting hyperplane of the sublevel set $\{y \in \mathbb{R}^d | f(y) \leq f(x)\}$ at $x$. For the special case $f = \gamma$, this fact is the basis for the concept of Birkhoff orthogonality, which we will discuss in Section 3.

**Example 2.1.** We have
\[
\partial_\varepsilon \gamma(x) = \begin{cases} \{x^* \in \mathbb{R}^d | \gamma^*(x^*) \leq 1\} = B(0, 1)^\circ, & x = 0, \\ \{x^* \in \mathbb{R}^d | \langle x^* | x \rangle \geq \gamma(x) - \varepsilon, \gamma^*(x^*) \leq 1\}, & \text{else}, \end{cases}
\tag{3}
\]
see [63, Theorem 2.4.2(ii)].
The first lemma serves as an \(\epsilon\)-version of Fermat’s rule [6, Theorem 16.2, Proposition 17.17] and can be proved analogously to its classical counterpart. Half of it is stated, e.g., in [30, Theorem XI.1.1.5].

**Lemma 2.2.** Let \(f : \mathbb{R}^d \to \mathbb{R}\) be a convex function and let \(x \in \mathbb{R}^d\), \(\epsilon \geq 0\). The following statements are equivalent:

(a) \(f(x) \leq f(y) + \epsilon\) for all \(y \in \mathbb{R}^d\),
(b) \(0 \in \partial f(x)\)
(c) \(f'(x; y) \geq 0\) for all \(y \in \mathbb{R}^d\).

**Proof.** (a)\(\iff\)(b): Both statements are equivalent to \(0 \leq f(y) - f(x) + \epsilon\) for all \(y \in \mathbb{R}^d\).
(b)\(\iff\)(c): We have

\[
0 \leq f(y) - f(x) + \epsilon \forall y \in \mathbb{R}^d \implies 0 \leq f(x + ty) - f(x) + \epsilon \forall y \in \mathbb{R}^d, t > 0
\]

\[
\implies 0 \leq \frac{f(x + ty) - f(x) + \epsilon}{t} \forall y \in \mathbb{R}^d, t > 0
\]

\[
\implies 0 \leq f'(x; y)
\]

(c)\(\iff\)(b):

\[
0 \leq f'(x; y) \forall y \in \mathbb{R}^d \implies 0 \leq f'(x; y - x) \forall y \in \mathbb{R}^d
\]

\[
\implies 0 \leq \frac{f(x + ty) - f(x) + \epsilon}{t} \forall y \in \mathbb{R}^d, t > 0
\]

\[
\implies 0 \leq f(y) - f(x) + \epsilon \forall y \in \mathbb{R}^d
\]

For \(\epsilon = 0\), Lemma 2.2 gives a characterization of minimizers of a convex function in terms of the subdifferential. The existence of a minimizer of a convex function can be guaranteed by the additional assumption of coercitivity, see [6, Theorem 11.9]. A function \(f : \mathbb{R}^d \to \mathbb{R}\) is called coercive if all of its sublevel sets \(\{x \in \mathbb{R}^d \mid f(x) \leq a\}\) are bounded.

The concept of approximate Birkhoff orthogonality is closely related to the following two proximality notions.

**Definition 2.3.** Let \(K\) be a non-empty closed convex set. A point \(x \in K\) is called an \(\epsilon\)-best approximation of \(y \in \mathbb{R}^d\) in \(K\) if \(\gamma(x - y) \leq \gamma(z - y) + \epsilon\) for all \(z \in K\). A point \(x \in K\) is called an \(\epsilon\)-best co-approximation of \(y \in \mathbb{R}^d\) in \(K\) if \(\gamma(x - z) \leq \gamma(y - z) + \epsilon\) for all \(z \in K\).

**Example 2.4.** For any \(y \in \mathbb{R}^d\), the set of \(\epsilon\)-best approximations of \(y\) in \(\mathbb{R}^d\) and the set of \(\epsilon\)-best co-approximations of \(y\) in \(\mathbb{R}^d\) coincide with \(B(y, \epsilon)\).

### 3 Approximate Birkhoff orthogonality

Approximate orthogonality relations are usually defined by introducing a relaxation parameter \(\epsilon\). For Birkhoff orthogonality in normed spaces, there is more than one approach...
to an approximate version. Some authors oaches of them are connected to semi-inner products [15,20] (see [41,47] for applications), but we will follow another approach [26,28] which works well with $\varepsilon$-subdifferentials and $\varepsilon$-best approximations.

**Definition 3.1.** The vector $x \in \mathbb{R}^d$ is called $\varepsilon$-Birkhoff orthogonal to $y \in \mathbb{R}^d$ (abbreviated by $x \perp^\varepsilon_B y$) if $\gamma(x) \leq \gamma(x + \lambda y) + \varepsilon$ for all $\lambda \in \mathbb{R}$. If $\varepsilon = 0$, we shall omit $\varepsilon$ from the notation and simply refer to Birkhoff orthogonality.

Trivially, the non-degeneracy property and the following homogeneity property are true for approximate Birkhoff orthogonality: For every $x, y \in \mathbb{R}^d$, $\lambda > 0$, and $\mu \in \mathbb{R}$, we have that $x \perp^\varepsilon_B y$ implies $\lambda x \perp^\varepsilon_B \mu y$. The remainder of this section is subdivided into three parts which address existence properties of $\varepsilon$-Birkhoff orthogonality, additivity of 0-Birkhoff orthogonality, and symmetry of $\varepsilon$-Birkhoff orthogonality, respectively.

### 3.1 Dual characterizations

First, we give a geometrical description of the $\varepsilon$-subdifferential of a gauge whose non-relaxed analog in normed spaces is [36, Theorem 2.1]. Note that given $x \in \mathbb{R}^d$ and $\varepsilon \geq \gamma(x)$, we have $x \perp^\varepsilon_B y$ for all $y \in \mathbb{R}^d$. Furthermore, (3) gives $\partial \gamma(x) = B(0,1)^\varepsilon$ independently of $\varepsilon$. Therefore, the restriction to $x \neq 0$ and $0 \leq \varepsilon \leq \gamma(x)$ is justified when asking for statements which link $\varepsilon$-subdifferentials to $\varepsilon$-Birkhoff orthogonality.

**Proposition 3.2.** Let $x, x^* \in \mathbb{R}^d$, $x \neq 0$, $\gamma^\circ(x^*) = 1$, and $0 \leq \varepsilon \leq \gamma(x)$. The following statements are equivalent:

(a) $x^* \in \partial^\varepsilon \gamma(x)$;
(b) $\langle x^* \mid x \rangle \geq \gamma(x) - \varepsilon$;
(c) $\langle x^* \mid x \rangle \geq 0, x \perp^\varepsilon_B h$ for all $h \in \mathbb{R}^d$ with $\langle x^* \mid h \rangle = 0$;
(d) $\sup \{ \langle x^* \mid z \rangle \mid z \in \mathbb{R}^d, \gamma(z) \leq \gamma(x) - \varepsilon \} \leq \langle x^* \mid x \rangle$.

**Proof.** (a)$\iff$(b): Since $x \neq 0$, one easily checks that

$$\partial^\varepsilon \gamma(x) = \left\{ x^* \in \mathbb{R}^d \mid \langle x^* \mid x \rangle \geq \gamma(x) - \varepsilon, \gamma^\circ(x^*) \leq 1 \right\},$$

see [63, Theorem 2.4.2(ii)].

(d)$\iff$(b): This follows from the identity

$$\sup_{z \in B(0,\gamma(x) - \varepsilon)} \langle x^* \mid z \rangle = \gamma^\circ(x^*)(\gamma(x) - \varepsilon) = \gamma(x) - \varepsilon,$$

which is a consequence of $\gamma^\circ = h_{B(0,1)}$.

(a)$\implies$(c): We have

$$\gamma(y) + \varepsilon \geq \gamma(x) + \langle x^* \mid y - x \rangle \quad \text{(4)}$$

for all $y \in \mathbb{R}^d$. If $y$ admits a representation $y = x + h$ with $h \in \mathbb{R}^d$, $\langle x^* \mid h \rangle = 0$, then (4) becomes

$$\gamma(x + h) + \varepsilon \geq \gamma(x) + \langle x^* \mid h \rangle = \gamma(x),$$

for all $y \in \mathbb{R}^d$. If $y$ admits a representation $y = x + h$ with $h \in \mathbb{R}^d$, $\langle x^* \mid h \rangle = 0$, then (4) becomes
that is, \( x \perp^\epsilon_B h \). By (b), we also have \( \langle x^* | x \rangle \geq \gamma(x) - \epsilon \geq 0 \).

(c)\(\implies\)(b): If \( x = 0 \), we are done. Otherwise, \( \langle x^* | x \rangle > 0 \). Thus \( \mu := \frac{\langle x^* | x \rangle}{\gamma(x) - \epsilon} > 0 \). For every \( y \in \mathbb{R}^d \) there are a number \( \lambda \in \mathbb{R} \) and a point \( h \in \mathbb{R}^d \) such that \( \langle x^* | h \rangle = 0 \) and \( y = \lambda x + h \).

If \( \lambda < 0 \), then \( \langle x^* | y \rangle = \lambda \langle x^* | x \rangle < 0 \leq \mu \gamma(y) \). If \( \lambda \geq 0 \), we obtain

\[
\langle x^* | y \rangle = \lambda \langle x^* | x \rangle = \lambda \mu (\gamma(x) - \epsilon) \leq \mu \gamma(\lambda x + h) = \mu \gamma(y).
\]

Thus \( 1 = \gamma^\epsilon(x^*) \leq \mu \) which is equivalent to \( \langle x^* | x \rangle \geq \gamma(x) - \epsilon \).

In case of \( \epsilon = 0 \), the previous result can be slightly improved to

**Corollary 3.3.** Let \( x, x^* \in \mathbb{R}^d \), \( x^* \neq 0 \). The following statements are equivalent:

(a) \( \frac{x^*}{\gamma^\epsilon(x^*)} \in \partial \gamma(x) \);

(b) \( \langle x^* | x \rangle = \gamma^\epsilon(x^*) \gamma(x) \);

(c) \( \langle x^* | x \rangle \geq 0 \), \( x \perp^\epsilon_B h \) for all \( h \in \mathbb{R}^d \) with \( \langle x^* | h \rangle = 0 \);

(d) \( \langle x^* | \cdot \rangle \) has a maximum on \( B(0, \gamma(x)) \) at \( x \).

The non-triviality of Proposition 3.2 is a consequence of the non-emptiness of the \( \epsilon \)-subdifferential of \( \gamma \), see [63, Theorem 2.4.9].

**Corollary 3.4** ([36, Theorem 2.2]). Let \( x \in \mathbb{R}^d \). Then there exists a vector \( x^* \in \mathbb{R}^d \setminus \{0\} \) such that \( x \perp^\epsilon_B h \) whenever \( \langle x^* | h \rangle = 0 \).

The next result links \( \epsilon \)-Birkhoff orthogonality and \( \epsilon \)-best approximations in linear subspaces to \( \epsilon \)-subdifferentials. The proof follows the lines of [59, Theorem 6.12] which is the corresponding result for normed spaces, see also [26, Lemma 1.1], [16, Theorem 2.5.1] and [60, Theorem 1.1, Theorem 5.1], [59, Lemma 1.1], [28, Theorem 2.3]

**Proposition 3.5.** Let \( U \subseteq \mathbb{R}^d \) be a non-trivial linear subspace, \( x, y \in \mathbb{R}^d \), \( y \notin U \), \( x \in U \) and \( \epsilon \geq 0 \). The following statements are equivalent:

(a) The point \( x \) is an \( \epsilon \)-best approximation of \( y \) in \( U \).

(b) We have \( (x - y) \perp^\epsilon U \) for all \( z \in U \).

(c) There exists \( x^* \in \mathbb{R}^d \) such that \( \gamma^\epsilon(x^*) = 1 \), \( \langle x^* | u \rangle = 0 \) for all \( u \in U \), and \( x^* \in \partial^\epsilon \gamma(x - y) \).

Moreover, the set of \( \epsilon \)-best approximations of \( y \) in \( U \) is non-empty, closed, and convex.

**Proof.** (a)\(\implies\)(b): Use \( x + \lambda U = U \).

(a)\(\implies\)(c): We have \( \rho := \inf_{z \in U} \gamma(x - y) > 0 \). Using the Hahn–Banach theorem, there exists a point \( x^* \in \mathbb{R}^d \), \( \gamma^\epsilon(x^*) = 1 \), \( \langle x^* | u \rangle = 0 \) for all \( u \in U \), and \( \langle x^* | x - y \rangle = \rho \). (Find a hyperplane \( H \) which contains \( U \) but not \( x - y \). Since every basis of \( H \) augmented by \( x - y \) forms a basis of \( \mathbb{R}^d \), the conditions \( \langle x^* | h \rangle = 0 \) for all \( h \in H \), \( \langle x^* | x - y \rangle = \rho \) uniquely determine \( x^* \), and \( \gamma^\epsilon(x^*) = 1 \) is satisfied.) Since \( x \) is an \( \epsilon \)-best approximation of \( y \) in \( U \), we have

\[
\gamma(x - y) \leq \rho + \epsilon = \langle x^* | x - y \rangle + \epsilon.
\]

(c)\(\implies\)(a): By virtue of (2) and (3), we have

\[
\gamma(x - y) \leq \langle x^* | x - y \rangle + \epsilon = \langle x^* | u - y \rangle + \epsilon \leq \gamma(u - y) + \epsilon
\]
for all \( u \in U \).

In order to prove the addendum, apply [6, Theorem 11.9] to the convex, hence continuous, and coercive function \( \gamma(-y) : U \to \mathbb{R} \) and conclude that the set of 0-best approximations of \( y \) in \( U \) is non-empty. Therefore, the set of \( \varepsilon \)-best approximations is also non-empty and, being a sublevel set of \( \gamma(-y) : U \to \mathbb{R} \), compact and convex.

Similarly, \( \varepsilon \)-best co-approximations of points in linear subspaces can be characterized using \( \varepsilon \)-Birkhoff orthogonality and \( \varepsilon \)-subdifferentials. Closely related results in finite-dimensional normed spaces are, for instance, [24, p. 1046, (1)], [50, Proposition 2.1], and [28, Theorems 2.3, 2.6, 2.10].

**Proposition 3.6.** Let \( U \subseteq \mathbb{R}^d \) be a non-trivial linear subspace, \( x, y \in \mathbb{R}^d, y \notin U, x \in U \) and \( \varepsilon \geq 0 \). The following statements are equivalent.

(a) The point \( x \) is an \( \varepsilon \)-best co-approximation of \( y \) in \( U \).

(b) We have \( z \perp^\varepsilon (y - x) \) for all \( z \in U \).

(c) For all \( z \in U \), \( x \) is an \( \varepsilon \)-best approximation of \( z \) in \( \langle x, y \rangle \).

(d) For \( z \in U \), there exists \( x^* \in \mathbb{R}^d \) such that \( \gamma(x^*) = 1, \langle x^* | u \rangle = 0 \) for all \( u \in U \), and \( x^* \in \partial_\varepsilon \gamma(y - x) \).

**Proof.** (a)\( \iff \) (b): Use \( x + \lambda U = U \).

(b)\( \implies \) (c): Plug \( x + U = U \) into (b) to obtain

\[
\gamma(x - z) \leq \gamma(y - z + \lambda(y - x)) + \varepsilon = \gamma(\lambda y + (1 - \lambda)x - z) + \varepsilon = \gamma(w - z) + \varepsilon
\]

for all \( z \in U, \lambda \in \mathbb{R} \), and \( w \in \langle x, y \rangle \).

(c)\( \implies \) (a): In

\[
\gamma(x - z) \leq \gamma(w - z) + \varepsilon \text{ for all } z \in U \text{ and } w \in \langle y, x \rangle ,
\]

we choose \( w = y \) to obtain

\[
\gamma(x - z) \leq \gamma(y - z) + \varepsilon \text{ for all } z \in U .
\]

(a)\( \iff \) (d): Combine the the methods for proving (a)\( \iff \) (c) in Proposition 3.5 and the equivalence (a)\( \iff \) (c).

As a corollary of Proposition 3.5 for one-dimensional subspaces, we obtain the following result, see also [36, Theorem 2.3].

**Corollary 3.7.** Let \( x, y \in \mathbb{R}^d, \varepsilon \geq 0 \), and \( a \in \mathbb{R} \). The following statements are equivalent:

(a) The point \( ax + y \) is an \( \varepsilon \)-best approximation of 0 in \( y + \text{lin}\{x\} \).

(b) The point \( ax \) is an \( \varepsilon \)-best approximation of \( -y \) in \( \text{lin}\{x\} \).

(c) We have \( (ax + y) \perp^\varepsilon x \).

(d) There exists \( x^* \in \mathbb{R}^d \) such that \( \gamma^\varepsilon(x^*) = 1, \langle x^* | x \rangle = 0 \) and \( \langle x^* | y \rangle \geq \gamma(ax + y) - \varepsilon \).

Moreover, the set of numbers \( a \in \mathbb{R} \) such that \( (ax + y) \perp^\varepsilon x \) is a compact interval.
The analogous results to [36, Corollary 2.2, Lemma 3.1] for the situation \( x \perp_B^\varepsilon (ax + y) \) are as follows.

**Proposition 3.8.** Let \( x, y \in \mathbb{R}^d \) and \( 0 \leq \varepsilon < \gamma(x) \). The set of numbers \( \alpha \in \mathbb{R} \) such that \( x \perp_B^\varepsilon (ax + y) \) is a compact interval. If \( x \neq 0 \) and \( x \perp_B^\varepsilon (ax + y) \), then \( |a| \leq \max \left\{ \frac{\gamma(y)}{\gamma(x) - \varepsilon}, \frac{\gamma(\varepsilon)}{\gamma(x) - \varepsilon} \right\} \).

**Proof.** By Corollary 3.4, there exists a vector \( x^* \in \mathbb{R}^d \setminus \{0\} \) such that \( x \perp_B^\varepsilon h \) whenever \( h \in \mathbb{R}^d \) and \( \langle x^* | h \rangle = 0 \). If \( x = 0 \), choose \( \alpha \in \mathbb{R} \) arbitrarily. Else, assume that, for all \( \alpha \in \mathbb{R} \), the vector \( x \) is not \( \varepsilon \)-Birkhoff orthogonal to \( ax + y \). That is, the line \( \{ax + y | \alpha \in \mathbb{R} \} \) does not intersect the hyperplane \( \{h \in \mathbb{R}^d | \langle x^* | h \rangle = 0 \} \). Consequently, \( \langle x^* | x \rangle = 0 \) and thus \( x = 0 \), a contradiction.

Now assume \( x \neq 0 \). If \( x \perp_B^\varepsilon (ax + y) \), then \( \gamma(x) - \varepsilon \leq \gamma(x + \lambda(ax + y)) \) for all \( \lambda \in \mathbb{R} \). In particular, for \( \lambda = -\frac{1}{\alpha} \), we obtain \( \gamma(x) - \varepsilon \leq \gamma\left(-\frac{1}{\alpha}y\right) \). If \( \alpha > 0 \), then \( \gamma(x) - \varepsilon \leq \frac{1}{\alpha} \gamma(-y) \). In case \( \alpha < 0 \) holds, we have \( \gamma(x) - \varepsilon \leq -\frac{1}{\alpha} \gamma(y) \). This yields \( |a| \leq \max \left\{ \frac{\gamma(y)}{\gamma(x) - \varepsilon}, \frac{\gamma(-y)}{\gamma(x) - \varepsilon} \right\} \) for \( \alpha \neq 0 \), which holds trivially for \( \alpha = 0 \).

Last, we prove that the intersection of \( G = \{z \in \mathbb{R}^d | x \perp_B^\varepsilon z \} \) with each line \( L \parallel \text{lin} \{x\} \) is a compact interval. If \( L = \{ax + y | \alpha \in \mathbb{R} \} \) is such a line, then \( G \cap L \subseteq \text{lin} \{x, y\} \). Hence we only have to consider the two-dimensional case, which is trivial. The set \( x + G \) is the union of all hyperplanes passing through \( x \) but not through the interior of \( B(0, \gamma(x) - \varepsilon) \).

Therefore, if we set

\[
C := \{\lambda u | \lambda \geq 0, u \in \text{lin} \{x, y\}, \gamma(u + x) \leq \gamma(x) - \varepsilon\},
\]

then \( C \) is a closed convex cone, we have \( C \cap (-C) = \{0\} \), and \( G = \text{cl}(\mathbb{R}^d \setminus (C \cup (-C))) \) is obviously directionally convex with respect to \( x \) which is an interior point of \( -C \), see Figure 1. \( \square \)

Generalizing [36, Theorem 3.2], we establish a necessary condition for \( \varepsilon \)-Birkhoff orthogonality in generalized Minkowski spaces, which is also a sufficient condition for \( \varepsilon = 0 \). The analogous statement for normed spaces and \( \alpha = 0 \) can be found in [29, p. 54].

**Theorem 3.9.** Let \( (\mathbb{R}^d, \gamma) \) be a generalized Minkowski space, \( x, y \in \mathbb{R}^d \), and \( \alpha \in \mathbb{R} \). Then \( x \perp_B^\varepsilon (ax + y) \) implies

\[
-\gamma'(x; -y) \leq -\alpha \gamma(x) \leq \gamma'(x; y).
\]  

The reverse implication is true for \( \varepsilon = 0 \).

**Proof.** Using the Cauchy–Schwarz inequality (2), we obtain the more general estimate

\[
\gamma'(x; ax + \mu y) = \max \{\alpha \langle x^* | x \rangle + \mu \langle x^* | y \rangle | x^* \in \partial_\varepsilon \gamma(x)\} \\
\leq \max \{\alpha \gamma(x) + \mu \langle x^* | y \rangle | x^* \in \partial_\varepsilon \gamma(x)\}
\]

\[
= \alpha \gamma(x) + \mu \max \{\langle x^* | y \rangle | x^* \in \partial_\varepsilon \gamma(x)\}
\]

\[
= \alpha \gamma(x) + \mu \gamma'(x; y)
\]

for all \( x, y \in \mathbb{R}^d, \alpha \in \mathbb{R}, \) and \( \mu \geq 0 \). Line (9) holds with equality for \( \varepsilon = 0 \). Now the claim follows from Lemma 2.2. \( \square \)
Figure 1. Proof of Proposition 3.8: The boundaries of the balls \( B(0, \gamma(x)) \) and \( B(0, \gamma(x) - \varepsilon) \) are depicted as bold lines. If \( C \) is the cone spanned by \( B(-x, \gamma(x) - \varepsilon) \), then the closure \( G \) of the complement of \( C \cup (-C) \) is the set of directions of straight lines through \( x \) which do not pass through the interior of \( B(0, \gamma(x) - \varepsilon) \).

Necessary and sufficient conditions for \( x \perp_B \varepsilon (ax + y) \) can be given in terms of linear functionals, see [36, Corollary 2.2] for the non-relaxed version in normed spaces.

Proposition 3.10. Let \( (\mathbb{R}^d, \gamma) \) be a generalized Minkowski space. Fix \( x, y \in \mathbb{R}^d \), \( a \in \mathbb{R} \), and define \( h : \mathbb{R} \to \mathbb{R} \), \( h(\lambda) = \gamma(x + \lambda(ax + y)) \). The following statements are true:

(a) \( x \perp_B (ax + y) \),
(b) \( g(0) \leq h(\lambda) + \varepsilon \) for all \( \lambda \in \mathbb{R} \),
(c) there exists \( x^* \in \mathbb{R}^d \) such that \( \gamma^o(x^*) = 1 \), \( \langle x^* | x \rangle \geq \gamma(x) - \varepsilon \) and \( \alpha = -\frac{\langle x^* | y \rangle}{\langle x^* | x \rangle} \),
(d) \( \gamma'_o(x; x^*; \pm(ax + y)) \geq 0 \).

Proof. The equivalence (a)\( \iff \) (b) is a direct consequence of the definition. The equivalences (b)\( \iff \) (c)\( \iff \) (d) follow from Lemma 2.2. This is because (c) and (d) are reformulations of the conditions \( 0 \in \partial_B h(0) \) and \( h'_o(0, \nu) \geq 0 \) for all \( \nu \in \mathbb{R} \), respectively.

Now [63, Theorem 2.8.10] yields

\[
\partial_x h(\lambda) = \bigcup_{\varepsilon_1, \varepsilon_2 \geq 0 \atop \varepsilon_1 + \varepsilon_2 = \varepsilon \atop q \in \partial_\gamma \gamma(x + \lambda(ax + y))} \partial_{\varepsilon_1} \langle x^* | x + (ax + y) \rangle(\lambda)
\]

\[
= \bigcup_{\varepsilon_1, \varepsilon_2 \geq 0 \atop \varepsilon_1 + \varepsilon_2 = \varepsilon} \{ \langle x^* | ax + y \rangle \}
\]

\[
x^* \in \partial_\gamma \gamma(x + \lambda(ax + y))
\]

\[
= \{ \langle x^* | ax + y \rangle \mid x^* \in \partial_\gamma \gamma(x + \lambda x) \}
\]

where the first equality holds because \( a : \mathbb{R} \to \mathbb{R} \), \( a(\lambda) = \langle x^* | x + \lambda(ax + y) \rangle \) is an affine function.
Using [63, Theorem 2.4.9], we obtain

\[
h'_s(\lambda, \nu) = \{\mu\nu \mid \mu \in \partial_{\epsilon} g(\lambda)\}
= \{(q \mid z) \mid q \in \partial_{\epsilon} \gamma(x + \lambda(ax + y))\}
= \{(q \mid vz) \mid q \in \partial_{\epsilon} \gamma(x + \lambda(ax + y))\}
= \gamma'(x + \lambda(ax + y), \nu(ax + y)).
\]

In particular,

\[
h'(0; \nu) = \gamma'(x; \nu(ax + y)), \quad \partial_{\epsilon} h(0) = \{(x^* \mid ax + y) \mid x^* \in \partial \gamma(x)\}.
\]

Taking the positive homogeneity of the \(\epsilon\)-directional derivative in the second variable into account, it suffices to consider \(\nu = \pm 1\) for checking whether \(h'(0; \nu) \geq 0\) for all \(\nu \in \mathbb{R}\). Moreover, we have

\[
0 \in \partial_{\epsilon} h(0) \iff \text{there exists } x^* \in \partial \gamma(x) \text{ such that } \langle x^* \mid ax + y \rangle = 0
\]
\[
\iff \text{there exists } x^* \in \mathbb{R}^d \text{ such that } \gamma'(x^*) \leq 1, \langle x^* \mid x \rangle \geq \gamma(x) - \epsilon, \alpha = \frac{\langle x^* \mid y \rangle}{\langle x^* \mid x \rangle}
\]
\[
\iff \text{there exists } x^* \in \mathbb{R}^d \text{ such that } \gamma'(x^*) = 1, \langle x^* \mid x \rangle \geq \gamma(x) - \epsilon, \alpha = \frac{\langle x^* \mid y \rangle}{\langle x^* \mid x \rangle}.
\]

This completes the proof. \(\square\)

Emulating key properties of usual inner products, semi-inner products enable Hilbert-like arguments in arbitrary Banach spaces. Prominent examples include the \textit{superior} and \textit{inferior semi-inner product} associated with a given norm \(\| \cdot \|\), which are in fact directional derivates of the convex function \(\| \cdot \|^2\). New approaches to classical concepts where developed, not only to optimization problems like the Fermat–Torricelli problem [17] and the best approximation problem [21], but also to geometric concepts like orthogonality [22, §§ 8–11]. In particular, several results connecting semi-inner products to Birkhoff orthogonality were derived. We close this subsection by demonstrating how Birkhoff orthogonality in generalized Minkowski spaces can be characterized in terms of natural analogs of the superior and inferior semi-inner products. To this end, consider the functions \(g : \mathbb{R}^d \rightarrow \mathbb{R}, g(x) = \frac{1}{2} \gamma(x)^2\), and \((\cdot, \cdot)_s, (\cdot, \cdot)_i : \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}\) defined by

\[
(y, x)_s = \lim_{\lambda \to 0} \frac{\gamma(x + \lambda y)^2 - \gamma(x)^2}{2\lambda} = g'(x; y),
\]
\[
(y, x)_i = \lim_{\lambda \to 0} \frac{\gamma(x + \lambda y)^2 - \gamma(x)^2}{2\lambda} = -g'(x; -y) = -(y, x)_s.
\]

Note that \((\cdot, \cdot)_s, (\cdot, \cdot)_i\) need not be semi-inner products in the sense of [22, Definition 6] since \((x, y)_p^p \leq (x, x)_p(y, y)_p^p\) for \(p \in (s, i)\) may be invalidated by the asymmetry of \(\gamma\). In normed
spaces, this estimate is checked in [22, Proposition 6]. The proof uses the reverse triangle inequality which is not valid for gauges. However, we can show an upper bound for \((\cdot, \cdot)_s:\)
\[
(y, x)_s = \lim_{\lambda \to 0} \frac{\gamma(x + \lambda y) - \gamma(x)}{\lambda}
\]
\[
= \lim_{\lambda \to 0} \frac{\gamma(x + \lambda y) + \gamma(x) - 2 \gamma(x)}{2 \lambda}
\]
\[
= \lim_{\lambda \to 0} \frac{\gamma(x + \lambda y) - \gamma(x)}{\lambda}
\]
\[
\leq \gamma(x) \lim_{\lambda \to 0} \frac{\gamma(\lambda y)}{\lambda}
\]
\[
\leq \gamma(x) \gamma(y),
\]
where line (10) can be written as \((y, x)_s = \gamma(x) \gamma'(x; y)\) which is basically the chain rule for directional derivatives [57, Proposition 3.6] applied to the function \(g\). Similarly, a lower bound for the function \((\cdot, \cdot)_i\) is \((y, x)_i = -\gamma(x) \gamma'(x; -y) \geq -\gamma(x) \gamma(-y)\).

Following the lines of [22, Proposition 5], we can also check that
\[
(x, x)_s = \lim_{\lambda \to 0} \frac{\gamma(x + \lambda x) - \gamma(x)}{\lambda}
\]
\[
= \gamma(x) \lim_{\lambda \to 0} \frac{(1 + \lambda)^2 \gamma(x) - \gamma(x)}{2 \lambda} = \gamma(x)^2
\]
and
\[
(x, x)_i = \lim_{\lambda \to 0, \lambda \to -1} \frac{\gamma(x + \lambda x) - \gamma(x)}{\lambda}
\]
\[
= \gamma(x) \lim_{\lambda \to 0, \lambda \to -1} \frac{(1 + \lambda)^2 \gamma(x) - \gamma(x)}{2 \lambda} = \gamma(x)^2
\]
for \(x \in X\).

For all \(x, y \in \mathbb{R}^d\), \(\alpha \in \mathbb{R}\), and \(\mu \geq 0\), a generalization of [22, Theorem 16] can be established by using the computation in the proof of Theorem 3.9 and the chain rule for directional derivatives
\[
g'(x; \alpha x + \mu y) = 2 \gamma(x) \gamma'(x; \alpha x + \mu y)
\]
\[
\leq 2 \gamma \left( \alpha \gamma(x) + \mu \gamma'(x; y) \right)
\]
\[
= 2 \alpha \gamma(x) + 2 \mu \gamma(x) \gamma'(x; y).
\]

In the classical theory in normed spaces, computations like these are the foundation for proving characterizations of Birkhoff orthogonality in terms of the superior and inferior semi-inner products. It turns out that such results can also be derived by using Theorem 3.9 for \(\epsilon = 0\). For instance, the analog of [22, Corollary 12] in generalized Minkowski spaces is the equivalence of the statements.
(a) \( x \perp_B (ax + y) \),
(b) \( (y, x)_i \leq -a\gamma(x)^2 \leq (y, x)_e \).

**Proof.** In (8), set \( \varepsilon = 0 \) and multiply by \( \gamma(x) \). \( \square \)

For \( \alpha = 0 \), this yields the equivalence of \( x \perp_B y \) and \( (y, x)_i \leq 0 \leq (y, x)_e \), see [22, Theorem 50] for the corresponding result in normed spaces.

### 3.2 Smoothness and rotundity

As in the case of normed spaces, Birkhoff orthogonality in generalized Minkowski spaces can be used to characterize rotundity and smoothness of the unit ball. The results for the general setting are the subject of this section. The first theorem connects smoothness and Birkhoff orthogonality; the corresponding results for normed spaces are [36, Theorems 4.2, 5.1].

**Theorem 3.11.** Let \((\mathbb{R}^d, \gamma)\) be a generalized Minkowski space. The following statements are equivalent:

(a) The unit ball \( B(0, 1) \) is smooth.
(b) The gauge \( \gamma \) is Gâteaux differentiable on \( \mathbb{R}^d \setminus \{0\} \).
(c) If \( x, y, z \in \mathbb{R}^d \), \( x \perp_B y \) and \( x \perp_B z \), then \( x \perp_B (y + z) \).
(d) For every \( x, y \in \mathbb{R}^d, x \neq 0 \), there exists a unique number \( \alpha \in \mathbb{R} \) such that \( x \perp_B (ax + y) \).
(e) For all \( x \in \mathbb{R}^d \) with \( \gamma(x) = 1 \), there exists a unique vector \( x^* \in \mathbb{R}^d \), \( \gamma(x^*) = 1 \) such that \( \langle x^* | x \rangle = 1 \).

In this case, \( x^* \) is the Gâteaux derivative of \( \gamma \) at \( x \) (items (e) and (b)), \( \gamma'(x; y) = -\alpha \gamma(x) \) (item (d)), and the unique hyperplane of \( B(0, 1) \) at one of its boundary points \( x \) consists of all points \( y \) such that \( x_0 \perp_B (y - x) \) (item (a)).

**Proof.** (e)\( \iff \) (a): See Proposition 3.2(b)\( \iff \) (d).

(a)\( \implies \) (c): The unique supporting hyperplane of \( B(0, \gamma(x)) \) passing through \( x \) has the form

\[
H := \left\{ h \in \mathbb{R}^d \left| \langle x^* | h \rangle = \langle x^* | x \rangle \right. \right\},
\]

where \( x^* \) is uniquely determined up to a constant factor. Using Proposition 3.2(c)\( \iff \) (d), we conclude that \( x \perp_B h \) if and only if \( h \in H - x \). That is, if \( x \perp_B y \) and \( x \perp_B z \), then \( y, z \in H - x \). Since \( H - x \) is a linear subspace of \( \mathbb{R}^d \), we also have \( y + z \in H - x \) and thus \( x \perp_B (y + z) \).

(c)\( \implies \) (d): Assume that there exist numbers \( \alpha, \beta \in \mathbb{R} \) such that \( x \perp_B (ax + y) \) and \( x \perp_B (\beta x + y) \). Due to the homogeneity of \( \perp_B \) and (c), we obtain \( x \perp_B (\alpha - \beta)x \) which implies \( \alpha = \beta \).

\( \neg \) (a)\( \implies \) \( \neg \) (d): Let \( x_1^*, x_2^* \in \mathbb{R}^d \) be vectors such that the hyperplanes

\[
H_i := \left\{ h \in \mathbb{R}^d \left| \langle x_i^* | h \rangle = \langle x_i^* | x \rangle \right. \right\}, \quad i \in \{1, 2\},
\]
are supporting hyperplanes of \( B(0, \gamma(x)) \) at \( x \). Set \( \alpha_i := -\langle x^* | x \rangle \). Then \( x \perp_B (\alpha_i x + y) \) for \( i \in \{1, 2\} \), and \( \alpha_1 \neq \alpha_2 \) if the intersection points of the line \( \{ax + y | a \in \mathbb{R} \} \) with \( H_1 \) and \( H_2 \) do not coincide.

(b) \( \implies \) (a): If \( \gamma \) has a Gâteaux derivate \( x^* \) at a point \( x \) with \( \gamma(x) = 1 \), then \( \partial \gamma(x) = \{x^*\} \). In particular, \( \gamma(x^*) = 1 \), and Proposition 3.2(a) \( \implies \) (d) yields the uniqueness of supporting hyperplanes of \( B(0, 1) \) at \( x \). Conversely, if there is a unique supporting hyperplane of \( B(0, 1) \) at \( x \), then Proposition 3.2(a) \( \implies \) (d) implies that \( \partial \gamma(x) \) is a singleton. Hence \( \gamma \) is Gâteaux differentiable at \( x \).

Alternatively, the implication (c) \( \implies \) (a) in Theorem 3.11 can be proved as follows. Let \( x \) be a boundary point of \( B(0, 1) \). The set \( \{z \in \mathbb{R}^d | x \perp_B z\} \) is a union of hyperplanes passing through the origin, hence a cone. If (c) is true, then this cone is convex. But a convex cone which is a union of hyperplanes is either a single hyperplane or \( \mathbb{R}^d \), the latter one contradicting \( x \neq 0 \).

The second theorem is a characterization of rotund gauges in terms of Birkhoff orthogonality; it generalizes [36, Theorems 4.3, 5.2].

**Theorem 3.12** ([36, Theorem 4.3, 5.2]). Let \((\mathbb{R}^d, \gamma)\) be a generalized Minkowski space. The following statements are equivalent:

(a) The unit ball \( B(0, 1) \) is rotund.
(b) The gauge \( \gamma \) is rotund.
(c) For every \( x, y \in \mathbb{R}^d \), \( x \neq 0 \), there exists a unique number \( \alpha \in \mathbb{R} \) such that \( \langle ax + y \rangle \perp_B x \).
(d) For all \( x^* \in \mathbb{R}^d \), the linear functional \( \langle x^* | \cdot \rangle \) has at most one maximum on \( B(0, 1) \).

**Proof.** (a) \( \implies \) (d): The family of supporting hyperplanes of \( B(0, 1) \) coincides with the family of hyperplanes

\[
\left\{ y \in \mathbb{R}^d \mid \langle x^* | y \rangle = h_B(x^*) \right\},
\]

where \( x^* \in \mathbb{R}^d \). The set of maximizers of \( \langle x^* | \cdot \rangle \) on \( B(0, 1) \) is then \( H \cap B(0, 1) \), which is a subset of the boundary of \( B(0, 1) \).

(a) \( \implies \) (b): See [32, Lemma 3.5].

(a) \( \implies \) (c): Fix \( x, y \in \mathbb{R}^d \), \( x \neq 0 \). Let \( \mu := \min \{\gamma(\lambda x + y) \mid \lambda \in \mathbb{R} \} \). Then \( B(0, \mu) \cap \{\lambda x + y \mid \lambda \in \mathbb{R} \} \) is the set of points \( \lambda x + y \) for which \( \lambda x + y \perp_B x \). It is a compact convex set, so if it is not a singleton, it is a segment, and there is a ball which contains a straight line segment in its boundary. Conversely, if \( B(0, 1) \) is not rotund, choose a segment \( [y, z] \subseteq \text{bd}(B(0, 1)) \) and set \( x = z - y \). Then the line \( \{\lambda x + y \mid \lambda \in \mathbb{R} \} \) does not meet the interior if \( B(0, 1) \) and, hence, \( \lambda x + y \perp_B x \) for \( \lambda \in (0, 1) \).

Theorem 3.11 shows that right additivity characterizes smoothness in all dimensions. However, left additivity does not play the same role for rotundity.

**Theorem 3.13.** Let \((\mathbb{R}^d, \gamma)\) be a generalized Minkowski space. If Birkhoff orthogonality is left additive, then \( \gamma \) is a norm.
Proof. Let $d = 2$. If $\gamma$ is not a norm, then there is an affine diameter not passing through the origin, see [61, § 4.1]. In other words, there are vectors $x, y, z \in \mathbb{R}^d$ such that $x$ and $y$ are linearly independent, $\gamma(x) = \gamma(y) = 1$, $x \not\perp_B z$, $y \perp_B z$, and $x - y \notin \text{lin}(z)$. Thus, there are numbers $\lambda, \mu \in \mathbb{R}$ with $0 < \lambda < 1$ and $\lambda x + (1 - \lambda)y = \mu z$. But left additivity and homogeneity imply $ax + by \perp_B z$ for rational numbers $a, b > 0$, and continuity of $\gamma$ yields the same thing for all numbers $a, b > 0$. Therefore $\mu z \perp_B z$, which implies $z = 0$.

If $d \geq 3$ and Birkhoff orthogonality is left additive, then it is left additive in each two-dimensional subspace of $\mathbb{R}^d$ which implies, using the first part of the proof, that the restriction of $\gamma$ to any two-dimensional subspace of $\mathbb{R}^d$ is a norm on that subspace. Hence, $\gamma$ itself is a norm.

Therefore, left-additivity for gauges reduces to the case of norms which can be found in [35, pp. 561, 562].

Corollary 3.14. Let $(\mathbb{R}^d, \gamma)$ be a generalized Minkowski space. Assume that Birkhoff orthogonality is left additive. If $d = 2$, then $\gamma$ is rotund. Else $\gamma$ is a norm induced by an inner product.

3.3 Orthogonality reversion and symmetry

Norms whose Birkhoff orthogonality relations coincide were studied in [55] and [51, Theorem 10], and the two-dimensional special case is implicitly stated, e.g., in [23, p. 165f.] and [62, p. 90]. The analogous investigation for gauges on $\mathbb{R}^2$ was done in [54, 4A], and it can be easily extended to the $d$-dimensional case.

The identification of pairs of norms whose Birkhoff orthogonality relations are inverses of each other yields the notion of antinorm in two-dimensional spaces, see [13, p. 867], [29, Proposition 3.1], and [43]. For normed spaces of dimension at least three, this class reduces to pairs of norms whose unit balls are homothetic ellipsoids, see [31, Theorem 3.2]. Closely related are norms whose Birkhoff orthogonality relation is symmetric. In the two-dimensional case, these norms are named after Johann Radon [52]. From [31, Theorem 3.2] it follows that in higher dimensions, symmetry of Birkhoff orthogonality characterizes Euclidean spaces. However, this result is much older and goes back to Blaschke [12].

In the present section, we prove that there are no asymmetric analogs of the antinorm and of Radon norms.

Theorem 3.15. If $\gamma_1, \gamma_2$ are gauges on $\mathbb{R}^d$ such that $x \perp_B y$ in $(\mathbb{R}^d, \gamma_1)$ if and only if $y \perp_B x$ in $(\mathbb{R}^d, \gamma_2)$, then $\gamma_1$ is a norm.

Proof. Let $x, y \in \mathbb{R}^d \setminus \{0\}$. Due to homogeneity and the assumption, we have

$$x \perp_B y \text{ in } (\mathbb{R}^d, \gamma_1) \iff y \perp_B x \text{ in } (\mathbb{R}^d, \gamma_2)$$

$$\iff y \perp_B \frac{-x}{\gamma_1(-x)} \text{ in } (\mathbb{R}^d, \gamma_2)$$

$$\iff \frac{-x}{\gamma(-x)} \perp_B y \text{ in } (\mathbb{R}^d, \gamma_1).$$

(11)
Case $\varepsilon = 0$: If $\gamma_1(x) = 1$, then $x$ and $\frac{x}{\gamma_1(x)}$ are the endpoints of a chord of the convex body $B := \{x \in \mathbb{R}^d \mid \gamma_1(x) \leq 1\}$ which passes through the origin 0. From (11) and the separation theorems it follows that there is a pair of parallel supporting hyperplanes of $B$ at the endpoints of every such chord. In other words, every chord passing through 0 is an affine diameter of $B$. Since 0 is an interior point of $B$, the claim follows by taking [61, § 4.1] into account.

Case $\varepsilon > 0$: Fix $x \in \mathbb{R}^d$ such that $\gamma_1(x) > \varepsilon$. Then there exists $y \in \mathbb{R}^d$ such that $x \perp_B y$ in $(\mathbb{R}^d, \gamma_1)$ and, without loss of generality, $\gamma_1(y) < \varepsilon$. But then $\gamma_1(y) < \varepsilon \leq \gamma_1(y + \lambda z) + \varepsilon$, so $y \perp_B z$ for all $z \in (\mathbb{R}^d, \gamma_1)$. By assumption, $z \perp_B y$ for all $z \in (\mathbb{R}^d, \gamma_2)$, that is,

$$\gamma_2(z) \leq \gamma_2(z + \lambda y) + \varepsilon$$

(12)

for all $z \in \mathbb{R}^d$ and $\lambda \in \mathbb{R}$. In particular, if we pick $n \in \mathbb{N}$ large enough such that $n\gamma(y) > \varepsilon$ and set $z = ny$ and $\lambda = -n$, then (12) becomes $n\gamma(y) \leq \varepsilon$, a contradiction. \qed

**Remark 3.16.** Birkhoff orthogonality in two-dimensional generalized Minkowski spaces can be “partially reversed” in the following sense which is patterned on the case of normed spaces, see again [13, p. 867]. Denote by $\rho : \mathbb{R}^2 \to \mathbb{R}^2$ the counterclockwise rotation by $90^\circ$ about the origin. Let $c : [0, 2\pi) \to \mathbb{R}^2$ be an injective parametrization of $S := \{x \in \mathbb{R}^d \mid \gamma(x) = 1\}$. Then, for all $t$, the directional derivative $c'(t; 1)$ exists (see [57, Proposition 3.6]), $c(t) \perp_B c'(t; 1)$ in $(\mathbb{R}^d, \gamma)$, and $c'(t; 1) \perp_B c(t)$ in $(\mathbb{R}^d, \gamma^\circ \circ \rho)$.

**Corollary 3.17.** Let $(\mathbb{R}^d, \gamma)$ be a generalized Minkowski space and let $\varepsilon \geq 0$. If $\varepsilon$-Birkhoff orthogonality is a symmetric relation, then $\varepsilon = 0$ and $\gamma$ is a norm.

## 4 Isosceles orthogonality

In Euclidean elementary geometry, triangles with reflection symmetry are isosceles, and if the lengths of the diagonals of a parallelogram coincide, the parallelogram is actually a rectangle. Formally, the orthogonality of vectors $x$ and $y$ in Euclidean space is equivalent to the equality of the lengths of the vectors $x + y$ and $x - y$. Serving as a definition in normed spaces, this yields the concept of *isosceles orthogonality*, which in general is different from Birkhoff orthogonality.

**Definition 4.1.** Let $(\mathbb{R}^d, \gamma)$ be a generalized Minkowski space. We say that the point $y \in \mathbb{R}^d$ is isosceles orthogonal to $x \in \mathbb{R}^d$ (abbreviated by $y \perp_I x$) if $\gamma(y + x) = \gamma(y - x)$.

### 4.1 Symmetry and directional convexity

In generalized Minkowski spaces, the isosceles orthogonality relation may be non-symmetric.

**Theorem 4.2.** If isosceles orthogonality is a symmetric relation, then $\gamma$ is a norm.

**Proof** (by J. Alonso). Take $u, v \in S(0, 1)$. Then $x = u + v$ is isosceles orthogonal to $y = u - v$ because $\gamma(x + y) = 2\gamma(u) = 2 = 2\gamma(v) = \gamma(x - y)$. If isosceles orthogonality is a symmetric relation, then also $2\gamma(-v) = \gamma(y - x) = \gamma(x + y) = 2$, that is, $-v \in S(0, 1)$. \qed
The set of points which are isosceles orthogonal to a given point \( x \in \mathbb{R}^d \) is also known as the *bisector* \( \text{bsc}(x, -x) \) of \( x \) and \(-x\), see [42, Definition 2.1.0.1]. The intersection of this bisector and every line parallel to \( \text{lin}(x) \) is non-empty, a fact which is stated for normed spaces, e.g., in [34, Theorem 4.4].

**Lemma 4.3.** Let \( x \) be a non-zero vector of a generalized Minkowski space \((\mathbb{R}^d, \gamma)\). Then, for all \( y \in \mathbb{R}^d \), there set of numbers \( \alpha \in \mathbb{R} \) with \((\alpha x + y) \perp I_x \) is a non-empty, closed, bounded, and convex, i.e., a compact interval.

**Proof.** For fixed \( y \in \mathbb{R}^d \), consider the function \( f : \mathbb{R} \to \mathbb{R} \),

\[
    f(\alpha) = \gamma(ax + y + x) - \gamma(ax + y - x).
\]

Note that \( f_{\alpha=0} = \{ \alpha \in \mathbb{R} \mid (ax + y) \perp I_x \} \). Closedness of this set is due to continuity of \( f \). For \( \alpha > 0 \), we have

\[
    \gamma((\alpha + \lambda)x + y)) - \gamma(ax + y) = \gamma\left(ax + \frac{\alpha}{\alpha + \lambda} y\right) - \gamma(ax + y) + \lambda \gamma\left(x + \frac{1}{\alpha + \lambda} y\right)
\]

provided \( \alpha + \lambda > 0 \). Using the subadditivity of \( \gamma \), we obtain

\[
    0 \leq \left| \gamma\left(ax + \frac{\alpha}{\alpha + \lambda} y\right) - \gamma(ax + y) \right| \leq \max\left\{ \gamma\left(\frac{\lambda}{\alpha + \lambda} y\right), \gamma\left(-\frac{\lambda}{\alpha + \lambda} y\right) \right\},
\]

yielding

\[
    \lim_{\alpha \to +\infty} \left( \gamma\left(ax + \frac{\alpha}{\alpha + \lambda} y\right) - \gamma(ax + y) \right) = 0.
\]

It follows that

\[
    \lim_{\alpha \to +\infty} \left( \gamma((\alpha + \lambda)x + y)) - \gamma(ax + y) \right) = \lim_{\alpha \to +\infty} \lambda \gamma\left(x + \frac{1}{\alpha + \lambda} y\right) = \lambda \gamma(x).
\]

Using this equation, we have

\[
    \begin{align*}
    \lim_{\alpha \to -\infty} (\gamma((ax + y) + x) - \gamma((ax + y) - x) & \\
& = \lim_{\alpha \to -\infty} (\gamma((\alpha + 1)x + y) - \gamma((\alpha - 1)x + y)) & \\
& = \lim_{\alpha \to -\infty} (\gamma((\alpha + 2)x + y) - \gamma(ax + y)) & \\
& = 2\gamma(x) > 0
    \end{align*}
\]

(13)

and

\[
    \begin{align*}
    \lim_{\alpha \to -\infty} (\gamma((ax + y) + x) - \gamma((ax + y) - x) & \\
& = \lim_{\alpha \to -\infty} (\gamma((-\alpha + 1)x + y) - \gamma((-\alpha - 1)x + y)) & \\
& = \lim_{\alpha \to -\infty} (\gamma((\alpha - 2)(-x) + y) - \gamma(a(-x) + y)) & \\
& = -2\gamma(-x) < 0.
    \end{align*}
\]

(14)
Using the intermediate value theorem, the continuity of \( f \) yields the existence of a zero of \( f \). Moreover, (13) and (14) imply that the set of zeros of \( f \) is bounded.

Now fix \( y \in X \) and \( \alpha > 0 \). We show that \( \gamma(y + x) - \gamma(y - x) \leq \gamma(ax + y + x) - \gamma(ax + y - x) \). If \( y \) is a multiple of \( x \), the claim is easily seen. Else the points \( y, -x, x \), and \( ax + y \) are (in this cyclic order) the vertices of a convex quadrangle. Now [33, Lemma 4.4] gives

\[
\gamma(y + x) + \gamma(ax + y - x) \leq \gamma(y - x) + \gamma(ax + y - x)
\]

or, equivalently,

\[
\gamma(y + x) - \gamma(y - x) \leq \gamma(ax + y + x) - \gamma(ax + y - x).
\]

Hence \( f \) is increasing, and its sublevel sets are intervals. This yields the convexity part of the claim. \( \square \)

In particular, Lemma shows that the intersection of \( \text{bsc}(-x, x) \) and every translate of \( \langle -x, x \rangle \) is either a singleton or a line segment. Next we will identify subsets of \( \text{bsc}(-x, x) \) which are unions of line segments parallel to \( \langle -x, x \rangle \). For \( p \in X \) and \( \phi \in X^* \), we define the set

\[
C(x, \phi) := x + \{ y \in X \mid \langle \phi \mid y \rangle = \gamma(y) \},
\]

which is the translation by \( p \) of the union of rays from the origin through the exposed face \( \phi = 1 \cap B(0, 1) \) of the unit ball \( B(0, 1) \).

**Proposition 4.4.** Given a non-zero vector \( x \) in a generalized Minkowski space \( (X, \gamma) \), we have \( C(-x, \phi) \cap C(x, \phi) \subseteq \text{bsc}(-x, x) \) whenever \( \phi \in X^* \) and \( \langle \phi \mid x \rangle = 0 \).

**Proof.** Consider

\[
z \in C(-x, \phi) \cap C(x, \phi)
\]

\[
\iff \begin{cases} 
\langle \phi \mid z + x \rangle = \gamma(z + x), \\
\langle \phi \mid z - x \rangle = \gamma(z - x)
\end{cases}
\]

\[
\implies \gamma(z + x) = \langle \phi \mid z + x \rangle + \langle \phi \mid -2x \rangle = \langle \phi \mid z - x \rangle = \gamma(z - x)
\]

\[
\implies z \in \text{bsc}(-x, x).
\]

\( \square \)

The cases in which the straight line \( y + \langle -x, x \rangle \) intersects the bisector \( \text{bsc}(-x, x) \) in at most one point are specified in Theorem 4.7 below. The following corollary of the triangle inequality serves as an auxiliary result.

**Lemma 4.5.** Let \( y, z \in (X, \gamma) \), \( \lambda \in (0, 1) \), and \( w = \lambda y + (1 - \lambda)z \). Then, for \( x \in X \),

\[
\gamma(w - x) \leq \max \{ \gamma(y - x), \gamma(z - x) \}
\]

with equality if and only if \( \gamma(w - x) = \gamma(y - x) = \gamma(z - x) \). In the case of equality, \( \gamma(w - x) = \min \{ \gamma(v - x) \mid v \in \langle y, z \rangle \} \) and \( \gamma(w - x) = \gamma(v - x) \) for all \( v \in \langle y, z \rangle \).
Proof. We have

\[
\gamma(w - x) = \gamma(\lambda y + (1 - \lambda)z - x)
\]
\[
= \gamma(\lambda(y - x) + (1 - \lambda)(z - x))
\]
\[
\leq \lambda \gamma(y - x) + (1 - \lambda)\gamma(z - x)
\]
\[
\leq \max \{\gamma(y - x), \gamma(z - x)\}
\] (15)

If (16) holds with equality, then \(\gamma(y - x) = \gamma(z - x)\), and if (15) holds with equality as well, then these numbers are equal to \(\gamma(w - x)\). In other words, \(y, w, \) and \(z\) are three collinear points on \(S(x, \gamma(w - x))\), hence \([y, z] \subseteq S(x, \gamma(w - x))\) or, equivalently, \(\gamma(w - x) = \gamma(v - x)\) for all \(v \in [y, z]\). Let \(p \in (y, z)\) such that \(z = \mu y + (1 - \mu)p\) for some \(\mu \in (0, 1)\). Applying the chain of inequalities above to \(y, z, \) and \(p,\) we obtain

\[
\gamma(z - x) \leq \max \{\gamma(y - x), \gamma(p - x)\}
\] (17)

Suppose \(\gamma(p - x) < \gamma(y - x)\). Then (17) holds with equality, i.e., \(\gamma(p - x) = \gamma(y - x)\). This is a contradiction. Thus \(\gamma(p - x) \geq \gamma(y - x)\), which shows \(\gamma(w - x) = \min \{\gamma(v - x) \mid v \in (y, z)\}\). 

Remark 4.6. Applying the above Lemma 4.5 to the generalized Minkowski space \((X, \gamma)\) where \(\tilde{\gamma}(x) := \gamma(-x)\), we get the same statements with reversed arguments, e.g., it holds \(\gamma(x - w) \leq \max \{\gamma(x - y), \gamma(x - z)\}\) for all \(x, y, z \in X\) and \(w \in [y, z]\).

Given a non-zero vector \(x\) in a generalized Minkowski space \((X, \gamma)\) with unit ball \(B = S(0, 1)\), we determine now for which vectors \(y \in X\) the intersection of the bisector \(bsc(-x, x)\) and the straight line \(y + (-x, x)\) is a singleton and for which \(y\) this intersection contains at least two points. If \(y\) and \(x\) are linearly depended, this intersection consists solely of the metric midpoint of \(-x\) and \(x\). Else the result depends on the shape of part of the unit sphere \(S(0, 1)\) lying in the two-dimensional half-flat \((-x, x) + [0, y]\). In particular, the ratio of the length of the segments \([-x, x]\) and the length of the maximal segment contained in \(S(0, 1) \cap ((-x, x) + [0, y])\) and parallel to \((-x, x)\) is important for the formulation of Theorem 4.7. Since we are taking ratios of lengths of parallel line segments, the result will not depend on whether we choose \(\gamma(z_1 - z_2), \gamma(z_2 - z_1),\) or \(2R([z_1, z_2], B)\) to be the (possibly oriented) length of the segment \([z_1, z_2]\) as long as we do it consistently. To show this, let \(w_1, w_2, z_1, z_2 \in X\) such that \(w_1 \neq w_2, z_1 \neq z_2\). Assume that there exist a number \(\lambda > 0\) such that \((z_1 - z_2) = \lambda(w_1 - w_2)\) or, equivalently, \((z_2 - z_1) = \lambda(w_2 - w_1)\). It follows that

\[
\frac{R([z_1, z_2], B)}{R([w_1, w_2], B)} = \frac{\gamma_B([z_1, z_2])}{\gamma_B([w_1, w_2])} = \lambda = \frac{\gamma_B(z_1 - z_2)}{\gamma_B(w_1 - w_2)} = \frac{\gamma_B(z_2 - z_1)}{\gamma_B(w_2 - w_1)}.
\]

Due to this, set \(hfl(x, y) = (-x, x) + [0, y]\) and

\[
M_y(x) := \sup \left\{\gamma_B(t - s) \mid [s, t] \subseteq S(0, 1) \cap hfl(x, y), \exists \lambda > 0 : t - s = \lambda x\right\}.
\]

for \(x, y \in X, x \neq 0\). Since the number \(M_y(x)\) only depends on \(hfl(x, y)\), the following generalization of [38, Theorem 2.6] is essentially a two-dimensional description of the bisector \(bsc(-x, x)\).
Theorem 4.7. Let $x$ and $y$ be non-zero vectors of a generalized Minkowski space $(X, \gamma)$. If $M_x(y) \leq \frac{2\gamma(y)}{\gamma(x)}$, then there exists a unique real number $\alpha$ such that $(y + \alpha x) \perp x$.

Proof. The existence of at least one number $\alpha \in \mathbb{R}$ with $y + \alpha x \in \text{bsc}(-x, x)$ follows from Lemma 4.1. Suppose that there are two numbers $\alpha_1, \alpha_2 \in \mathbb{R}$ such that $\alpha_1 < \alpha_2$ and $(y + \alpha_1 x, y + \alpha_2 x) \subseteq \text{bsc}(-x, x)$. Let $f : \mathbb{R} \to \mathbb{R}$, $f(\lambda) = \gamma(y + \lambda x)$. We have

$$f(\alpha_1 + 1) = \gamma(y + \alpha_1 x + x) = \gamma(y + \alpha_1 x - x) = f(\alpha_1 + 1)$$
$$f(\alpha_2 + 1) = \gamma(y + \alpha_2 x + x) = \gamma(y + \alpha_2 x - x) = f(\alpha_2 + 1)$$

Since $f$ is a convex function, equations (18) and (19) imply that $f$ is constant on $[\alpha_1 - 1, \alpha_2 + 1]$. By Lemma 4.5, this constant equals $\eta := \min \{f(\lambda) | \lambda \in \mathbb{R}\}$. Therefore, the line segment $[y + (\alpha_1 - 1)x, y + (\alpha_2 + 1)x]$ is contained in $S(0, \eta)$ and we have

$$M_x(y) \geq \frac{1}{\eta} (\alpha_2 - \alpha_1 + 2) \gamma(x) \geq \frac{1}{\gamma(y)} (\alpha_2 - \alpha_1 + 2) \gamma(x) > 2 \frac{\gamma(x)}{\gamma(y)}$$

An intriguing and, surprisingly, characteristic property of bisectors in the Euclidean space is their hyperplanarity, see [33, Proposition 4.10]. Closely related

4.2 Characterizations of norms

Theorem 4.8. Let $(X, \gamma)$ be a generalized Minkowski space. The following statements are equivalent.

(a) The gauge $\gamma$ is a norm induced by an inner product.
(b) Isosceles orthogonality is right homogeneous.
(c) Isosceles orthogonality is right additive.
(d) Isosceles orthogonality is left homogeneous.
(e) Isosceles orthogonality is left additive.

Proof. It is sufficient to show that right additivity, left additivity, right homogeneity, and left homogeneity imply that $\gamma$ is a norm. The claim then follows from [34, Theorem 4.7, Theorem 4.8].

(b): If $y \perp x$, then $y \perp \lambda x$ for all $\lambda > 0$ or, equivalently, $\lambda^{-1} y \in \text{bsc}(x, -x)$ for all $\lambda > 0$. Taking the limit $\lambda \to +\infty$, we obtain $0 \in \text{bsc}(-x, x)$. (Note that the bisector is a closed set by continuity of $\gamma$.) Since $x$ was chosen arbitrarily, $\gamma$ is a norm.

(c): Like before, but with $\lambda \in \mathbb{N}$ instead of $\lambda > 0$.

(d): Given $x \in \mathbb{R}^d \setminus \{0\}$, there exists exactly one number $\alpha$ for which $ax \perp x$. If $\gamma$ is not a norm, then $x$ can be chosen such that $\alpha \neq 0$. By left additivity, $\lambda ax \perp x$ for all $\lambda > 0$ which is impossible for $|\lambda a| > 1$.

(e): Like before, but with $\lambda \in \mathbb{N}$ instead of $\lambda > 0$. 

19
Within the class of normed spaces, Birkhoff orthogonality implies isosceles orthogonality if and only if the space is a Hilbert space, see [48, Theorem 2], [4, (10.2)]. A complementary statement is valid for normed spaces: Isosceles orthogonality implies Birkhoff orthogonality if and only if the space is a Hilbert space, see [48, Theorem 1] and [4, (10.9)]. The same is true for generalized Minkowski spaces.

**Theorem 4.9.** Let $(\mathbb{R}^d, \gamma)$ be a generalized Minkowski space.

(a) If Birkhoff orthogonality implies isosceles orthogonality, then $\gamma$ is a norm.
(b) If isosceles orthogonality implies Birkhoff orthogonality, then $\gamma$ is a norm.

**Proof.** (a): We have $0 \perp_B y$ for all $y \in \mathbb{R}^d$, thus $\gamma(y) = \gamma(-y)$ for all $y \in \mathbb{R}^d$.

(b): Assume that $\gamma$ is not a norm. Then there exists $y \in \mathbb{R}^d$ such that $\gamma(y) \neq \gamma(-y)$. Furthermore, there is a unique point $x \in (-y,y)$ such that $x \perp_I y$, namely $x = \frac{\gamma(-y) - \gamma(y)}{\gamma(-y) - \gamma(y)} y \neq 0$.

Due to the hypothesis, we have $x \perp_B \frac{\gamma(-y) + \gamma(y)}{\gamma(-y) - \gamma(y)} x$, which is impossible.

### 5 Final remarks

For extending the notion of orthogonality from Euclidean space to arbitrary normed spaces, there are various alternatives each of which has its own benefits (see [2, 3] for an overview of orthogonality types in normed spaces). By replacing the norm by a gauge, we translated two of the notions from normed spaces to generalized Minkowski spaces. Apart from this extension of the geometric setting, the relaxation of the orthogonality relation itself has been approached not only in the way presented here but also differently. Dragomir [20] introduced an approximate Birkhoff orthogonality relation $x \perp y$ via $\|x \pm \lambda y\| \leq \|x\| + \epsilon \|x\|$ for all $\lambda \in \mathbb{R}$. This condition is a left-homogeneous version of the one acting as the model for Definition 3.1. Chmieliński [15] discussed two approximate orthogonality relations in normed spaces defined via $\|x + \lambda y\|^2 \geq \|x\|^2 - 2\varepsilon \langle x, \lambda y \rangle$ and $\|x + \lambda y\| \geq \sqrt{1 - \varepsilon^2} \|x\|$ (for all $\lambda \in \mathbb{R}$ in each case), respectively, the latter one being a reparametrization of Dragomir’s condition. Both relations are left-homogeneous and right-homogeneous as well. In inner-product spaces, the condition $\|x + \lambda y\| \geq \sqrt{1 - \varepsilon^2} \|x\|$ for all $\lambda \in \mathbb{R}$ is equivalent to $|\langle y, x \rangle| \leq \varepsilon \|x\| \|y\|$. Due to the close relationship between orthogonality and the Cauchy–Schwarz inequality (see Corollary 3.3), the relaxed inequality $|\langle y, x \rangle| \leq \varepsilon \|x\| \|y\|$ itself might serve as a definition of approximate orthogonality. Here, 0-orthogonality is the Euclidean orthogonality (independently of the chosen norm), 1-orthogonality holds trivially because of the Cauchy–Schwarz inequality. (So the interesting cases which reflect the geometry of the normed space will have $0 < \varepsilon < 1$.) However, since $|\langle y, x \rangle| \leq \gamma(x)\gamma(y)$ is wrong in general, relaxing the Cauchy–Schwarz inequality in a multiplicative way has to be done differently in generalized Minkowski spaces. Chmieliński [15] and Dragomir [22, §§ 8-11] linked (relaxed) Birkhoff-type orthogonality notions to semi-inner products, not only the superior and inferior semi-inner products whose gauge counterparts appear in the end of Subsection 3.1. Can one nicely extend the Dragomir–Chmieliński definitions to generalized Minkowski spaces in order to obtain similar results? What is a suitable substitute for semi-inner products?
In the interplay between orthogonality types, metric projections onto linear subspaces, the radial projection onto the unit ball, and, of course, related characterizations of special classes of Banach spaces, also several constants and moduli which describe the geometry of the underlying space take part. Notable examples are the James constant [27, 45], the Dunkl-Williams constant [46], the rectangular constant [5, 7, 25, 39, 40], and the Schäffer–Thele constant which also coincides with the bias and the metric projection bounds of Smith, Baronti, and Franchetti [18]. (Note that the rectangular constant and the Schäffer–Thele constant are special values of the rectangular modulus introduced in [56].) To our best knowledge, such notions have not been investigated in generalized Minkowski spaces.

In normed spaces, isosceles orthogonality is trivially a symmetric relation. This is not the case for all other gauges. In view of Lemma 4.1, the following question has to be answered separately: For given vectors \( x, y \in \mathbb{R}^d \), \( x \neq 0 \), in a generalized Minkowski space \((\mathbb{R}^d, \gamma)\), is there a number \( a \in \mathbb{R} \) such that \( x \perp_{I} (ax + y) \)?

References

[1] Y. I. Alber, *James orthogonality and orthogonal decompositions of Banach spaces*, J. Math. Anal. Appl. 312 (2005), no. 1, pp. 330–342, doi: 10.1016/j.jmaa.2005.03.027.

[2] J. Alonso and C. Benítez, *Orthogonality in normed linear spaces: a survey. I. Main properties*, Extracta Math. 3 (1988), no. 1, pp. 1–15.

[3] ______, *Orthogonality in normed linear spaces: a survey. II. Relations between main orthogonalities*, Extracta Math. 4 (1989), no. 3, pp. 121–131.

[4] D. Amir, *Characterizations of Inner Product Spaces*, Operator Theory: Advances and Applications, vol. 20, Birkhäuser Verlag, Basel, 1986, doi: 10.1007/978-3-0348-5487-0.

[5] M. Baronti, *Su alcuni parametri degli spazi normati*, Boll. Unione Mat. Ital., V. Ser., B 18 (1981), pp. 1065–1085.

[6] H. H. Bauschke and P. L. Combettes, *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*, CMS Books in Mathematics, Springer, New York, 2011, doi: 10.1007/978-1-4419-9467-7.

[7] C. Benítez and M. del Río, *The rectangular constant for two-dimensional spaces*, J. Approximation Theory 19 (1977), no. 1, pp. 15–21.

[8] C. Benítez, M. Fernández, and M. L. Soriano, *Orthogonality of matrices*, Linear Algebra Appl. 422 (2007), no. 1, pp. 155–163, doi: 10.1016/j.laa.2006.09.018.

[9] T. Bhattacharyya and P. Grover, *Characterization of Birkhoff–James orthogonality*, J. Math. Anal. Appl. 407 (2013), no. 2, pp. 350–358, doi: 10.1016/j.jmaa.2013.05.022.
[10] G. Birkhoff, *Orthogonality in linear metric spaces*, Duke Math. J. 1 (1935), no. 2, pp. 169–172, doi: 10.1215/S0012-7094-35-00115-6.

[11] P. V. M. Blagojević and R. N. Karasev, *The Schwarz genus of the Stiefel manifold*, Topology Appl. 160 (2013), no. 18, pp. 2340–2350, doi: 10.1016/j.topol.2013.07.028.

[12] W. Blaschke, *Räumliche Variationsprobleme mit symmetrischer Transversalitätsbedingung*, Leipz. Ber. 68 (1916), pp. 50–55.

[13] H. Busemann, *The isoperimetric problem in the Minkowski plane*, Amer. J. Math. 69 (1947), pp. 863–871.

[14] Z.-Z. Chen, W. Lin, and L.-L. Luo, *Projections, Birkhoff orthogonality and angles in normed spaces*, Commun. Math. Res. 27 (2011), no. 4, pp. 378–384.

[15] J. Chmieliński, *On an ε-Birkhoff orthogonality*, J. Inequal. Pure Appl. Math. 6 (2005), no. 3, 7 pp.

[16] S. Cobzas, *Functional Analysis in Asymmetric Normed Spaces*, Frontiers in Mathematics, Birkhäuser/Springer, Basel, 2013, doi: 10.1007/978-3-0348-0478-3.

[17] D Comănescu, S. S. Dragomir, and E. Kikianty, *Toricellian points in normed linear spaces*, J. Inequal. Appl. (2013), 15 pp., doi: 10.1186/1029-242X-2013-258.

[18] J. Desbiens, *Sur les constantes de Thele et de Schäffer*, Ann. Sci. Math. Québec 16 (1992), no. 2, pp. 125–141.

[19] M. Deza and M. D. Sikirić, *Voronoi polytopes for polyhedral norms on lattices*, Discrete Appl. Math. 197 (2015), pp. 42–52, doi: 10.1016/j.dam.2014.09.007.

[20] S. S. Dragomir, *On approximation of continuous linear functionals in normed linear spaces*, An. Univ. Timişoara Ser. Științ. Mat. 29 (1991), no. 1, pp. 51–58.

[21] , *Characterization of best approximants from level sets of convex functions in normed linear spaces*, Nonlinear Funct. Anal. Appl. 6 (2001), no. 1, pp. 89–93.

[22] , *Semi-inner products and applications*, Nova Science Publishers, Inc., Hauppauge, NY, 2004.

[23] H. G. Eggleston, *Sets of constant width in finite dimensional Banach spaces*, Israel J. Math. 3 (1965), no. 3, pp. 163–172, doi: 10.1007/BF02759749.

[24] C. Franchetti and M. Furi, *Some characteristic properties of real Hilbert spaces*, Rev. Roumaine Math. Pures. Appl. 17 (1972), pp. 1045–1048.

[25] P. Ghosh, D. Sain, and K. Paul, *On rectangular constant in normed linear spaces*, 2014, arxiv: 1407.1353v1.
[26] M. R. Haddadi, H. Mazaheri, and A. Shoja, ε-general orthogonality and ε-orthogonality in normed linear spaces, J. Math. Sci. Adv. Appl. 3 (2009), no. 1, pp. 71–75.

[27] C. Hao and S. Wu, Homogeneity of isosceles orthogonality and related inequalities, J. Inequal. Appl. (2011), p. 9 pp., doi: 10.1186/1029-242X-2011-84.

[28] R. Hasani, H. Mazaheri, and S. M. Vaezpour, On the ε-best coapproximation, Int. Math. Forum 2 (2007), no. 53-56, pp. 2599–2606.

[29] L. Hetzelt, On suns and cosuns in finite-dimensional normed real vector spaces, Acta Math. Hungar. 45 (1985), no. 1-2, pp. 53–68.

[30] J.-B. Hiriart-Urruty and C. Lemaréchal, Convex Analysis and Minimization Algorithms. II, Grundlehren der Mathematischen Wissenschaften, vol. 306, Springer, Berlin, 1993, doi: 10.1007/978-3-662-06409-2.

[31] Á. G. Horváth, Zs. Lángi, and M. Spirova, Semi-inner products and the concept of semi-polarity, Results Math. (2015), p. 18 pp., doi: 10.1007/s00025-015-0510-y.

[32] T. Jahn, Y. S. Kupitz, H. Martini, and C. Richter, Minsum location extended to gauges and to convex sets, J. Optim. Theory Appl. 166 (2015), no. 3, pp. 711–746, doi: 10.1007/s10957-014-0692-6.

[33] T. Jahn, H. Martini, and C. Richter, Bi- and multifocal curves and surfaces for gauges, J. Convex Anal. 23 (2016), no. 3, 40 pp.

[34] R. C. James, Orthogonality in normed linear spaces, Duke Math. J. 12 (1945), pp. 291–302.

[35] ———, Inner product in normed linear spaces, Bull. Amer. Math. Soc. 53 (1947), pp. 559–566, doi: 10.1090/S0002-9904-1947-08831-5.

[36] ———, Orthogonality and linear functionals in normed linear spaces, Trans. Amer. Math. Soc. 61 (1947), pp. 265–292, doi: 10.2307/1990220.

[37] K. Jha, K. Paul, and D. Sain, On strong orthogonality and strictly convex normed linear spaces, J. Inequal. Appl. (2013), 7 pp., doi: 10.1186/1029-242X-2013-242.

[38] D. Ji, J. Li, and S. Wu, On the uniqueness of isosceles orthogonality in normed linear spaces, Results Math. 59 (2011), no. 1-2, pp. 157–162, doi: 10.1007/s00025-010-0069-6.

[39] J. L. Joly, Caractérisations d’espaces hilbertiens au moyen de la constante rectangle, J. Approximation Theory 2 (1969), pp. 301–311.

[40] O. P. Kapoor and S. B. Mathur, Some geometric characterizations of inner product spaces, Bull. Austral. Math. Soc. 24 (1981), no. 2, pp. 239–246, doi: 10.1017/S0004972700007619.
[41] M. Karamanlis and P. J. Psarrakos, *Birkhoff–James ε-orthogonality sets in normed linear spaces*, The Natália Bebiano Anniversary Volume, Textos Mat. Sér. B, vol. 44, Univ. Coimbra, Coimbra, 2013, pp. 81–92.

[42] L. Ma, *Bisectors and Voronoi Diagrams for Convex Distance Functions*, Ph.D. thesis, Fernuniversität Hagen, Hagen, 2000.

[43] H. Martini and K. Swanepoel, *The geometry of Minkowski spaces—a survey, Part II*, Expo. Math. 22 (2004), no. 2, pp. 93–144, doi: 10.1016/S0723-0869(04)80009-4.

[44] H. Martini and S. Wu, *On Zindler curves in normed planes*, Canad. Math. Bull. 55 (2012), no. 4, pp. 767–773, doi: 10.4153/CMB-2011-112-x.

[45] ______, *Orthogonalities, transitivity of norms and characterizations of Hilbert spaces*, Rocky Mountain J. Math. 45 (2015), no. 1, pp. 287–301, doi: 10.1216/RMJ-2015-45-1-287.

[46] H. Mizuguchi, *Geometrical Constants and Norm Inequalities in Banach Spaces*, Ph.D. thesis, Niigata University, Niigata, 2013, http://hdl.handle.net/10191/24070.

[47] M. S. Moslehian and A. Zamani, *Approximate Roberts orthogonality*, Aequationes Math. 89 (2015), no. 3, pp. 529–541, doi: 10.1007/s00010-013-0233-7.

[48] K. Ôhira, *On some characterizations of abstract Euclidean spaces by properties of orthogonality*, Kumamoto J. Sci. Ser. A. 1 (1952), no. 1, pp. 23–26.

[49] W. Pacheco Redondo and T. Rosas Soto, *On orthocentric systems in Minkowski planes*, Beitr. Algebra Geom. 56 (2015), no. 1, pp. 249–262, doi: 10.1007/s13366-014-0214-6.

[50] P. L. Papini and I. Singer, *Best coapproximation in normed linear spaces*, Monatsh. Math. 88 (1979), no. 1, pp. 27–44, doi: 10.1007/BF01305855.

[51] T. Precupanu, *Duality mapping and Birkhoff orthogonality*, An. Stiint. Univ. Al. I. Cuza Iaşi. Mat. (S.N.) 59 (2013), no. 1, pp. 103–112.

[52] J. Radon, *Über eine besondere Art ebener konvexer Kurven*, Leipz. Ber. 68 (1916), pp. 123–128.

[53] G. Samorodnitsky and M. S. Taqqu, *Stable Non-Gaussian Random Processes: Stochastic Models With Infinite Variance*, Stochastic Modeling, Chapman & Hall, New York, 1994.

[54] J. J. Schäffer, *Geometry of Spheres in Normed Spaces*, Lecture Notes in Pure and Applied Mathematics, vol. 20, Marcel Dekker, Inc., New York-Basel, 1976.

[55] P. Schöpf, *Orthogonality and proportional norms*, Anz. Österreich. Akad. Wiss. Math.-Natur. Kl. 133 (1996), pp. 11–16.

[56] I. Serb, *Rectangular modulus, Birkhoff orthogonality and characterizations of inner product spaces*, Comment. Math. Univ. Carolin. 40 (1999), no. 1, pp. 107–119.
[57] A. Shapiro, *On concepts of directional differentiability*, J. Optim. Theory Appl. 66 (1990), no. 3, pp. 477–487, doi: 10.1007/BF00940933.

[58] E. N. Shonoda and N. Omar, *Some special curves in normed three-dimensional space*, Appl. Math. Inf. Sci. Lett. 2 (2014), no. 2, pp. 53–57, doi: 10.12785/amisl/020204.

[59] I. Singer, *Best Approximation in Normed Linear Spaces by Elements of Linear Subspaces*, Die Grundlehren der mathematischen Wissenschaften, vol. 171, Springer, New York/Berlin, 1970.

[60] _____, *The Theory of Best Approximation and Functional Analysis*, Society for Industrial and Applied Mathematics, Philadelphia, PA, 1974.

[61] V. Soltan, *Affine diameters of convex bodies—a survey*, Expo. Math. 23 (2005), no. 1, pp. 47–63, doi: 10.1016/j.exmath.2005.01.019.

[62] A. C. Thompson, *Minkowski Geometry*, Cambridge University Press, Cambridge, 1996.

[63] C. Zălinescu, *Convex Analysis in General Vector Spaces*, World Scientific Publishing, River Edge, NJ, 2002, doi: 10.1142/9789812777096.