Canad. Math. Bull. Vol. 64 (4), 2021, pp. 828–839
http://dx.doi.org/10.4153/S0008439520000867

Surjective isometries of metric geometries

A. F. Beardon and D. Minda

Abstract. Many authors define an isometry of a metric space to be a distance-preserving map of the space onto itself. In this note, we discuss spaces for which surjectivity is a consequence of the distance-preserving property rather than an initial assumption. These spaces include, for example, the three classical (Euclidean, spherical, and hyperbolic) geometries of constant curvature that are usually discussed independently of each other. In this partly expository paper, we explore basic ideas about the isometries of a metric space, and apply these to various familiar metric geometries.

1 Introduction

The paper [28] begins with the sentence "One of the basic problems in geometric investigations is this: given a space S endowed with a metric d, describe the group of isometries of S with respect to the metric d." This sentence describes the content of this paper, although we will take a very different path to that taken in [28]. A map $f: X \to X$ is an *isometry* of a metric space (X, d) if it preserves distances; that is, if, for all x and y in X, we have $d\big(f(x), f(y)\big) = d(x, y)$. Obviously, an isometry is injective. Although many authors find it convenient to include surjectivity in their definition of an isometry, we will not do this; indeed, the motivation behind this paper was to examine circumstances under which an isometry must necessarily be surjective. Here, we will follow the terminology of Busemann [9, p. 15] and use the word *motion* to signify a *surjective isometry*. Since any injective map (whether surjective or not) of a set into itself is an isometry with respect to the discrete metric, there are many metric spaces that support both motions (surjective isometries) and non-surjective isometries. However, in all cases, the set of motions is a group under composition. We denote the group of motions by $\mathrm{M}(X)$, and the semigroup of isometries by $\mathrm{I}(X)$. Throughout, if f and g are maps, then fg (if defined) denotes the composite map $x \mapsto f(g(x))$, and f^n denotes the n-th iterate of f.

The second half of the paper concerns the isometry groups for Minkowski spaces that are finite-dimensional, real, normed vector spaces. The known isomorphism groups for the ℓ^p norms on $\mathbb{R}^n$ are established in a simple manner. For $n = 2$, we find the isometry group for norms with unit ball equal to a regular $2k$-gon for any positive integer k. There does not exist a norm on $\mathbb{R}^n$ with unit ball equal to a closed regular polygon with an odd number of sides.

Received by the editors May 15, 2020; revised September 23, 2020.
Published online on Cambridge Core October 28, 2020.
AMS subject classification: **53A30**, 53A35, 30F99.
Keywords: Constant curvature geometries, isometries.

2 Isometries and Motions of a Metric Space

The following sufficient condition for an isometry to be surjective is well known, and, for completeness, we give the short proof.

Lemma 2.1 *Each isometry f of a compact metric space (X, d) is surjective.*

Proof Choose any y in X. By compactness, the sequence $y, f(y), f^2(y), \dots$ has a convergent subsequence so, for any positive ε, there are positive integers p and q such that $d(f^p(y), f^{p+q}(y)) < \varepsilon$. As $d(y, f^q(y)) = d(f^p(y), f^{p+q}(y)) < \varepsilon$, this implies that y is in the closure of $f(X)$. As $f(X)$ is compact, it is closed; thus, $y \in f(X)$. ■

Lemma 2.1 has the following consequence.

Corollary 2.2 *Suppose that X and Y are metric spaces, that X is compact, and that there is an isometry g of Y into X. Then any isometry $f: X \to Y$ is surjective, so Y is isometrically equivalent to X.*

Proof The map gf is an isometry of X into itself so, from Lemma 2.1, $gf(X) = X$. Thus, $g(Y) \subset X = gf(X) \subset g(Y)$, which shows that $g(Y) = X$. Thus, g is a bijection from Y to X so that $Y = g^{-1}(X) = f(X)$ and Y is isometrically equivalent to X. ■

We can weaken the assumption of compactness in Lemma 2.1 at the cost of including the existence of a fixed point. First, a metric space is said to be *finitely compact* if each closed ball is compact.

Lemma 2.3 *If an isometry f of a finitely compact metric space (X, d) has a fixed point, then it is surjective.*

Proof Suppose that f has a fixed point y, and that $r > 0$, and let B_r be the closed (compact) ball with centre y and radius r. Then f is an isometry of B_r into itself, so, by Lemma 2.1, $f(B_r) = B_r$. As $X = \cup_{r>0} B_r$, this implies that $f(X) = X$. ■

Next, we have a simple result that guarantees that *all isometries are motions*. The set $\mathrm{I}(X)$ of isometries of X *acts transitively* on X if, given any a and b in X, there is an isometry g such that $g(a) = b$. Note that g need not be a motion, and that $\mathrm{I}(X)$ acts transitively if merely some set of isometries acts transitively.

Theorem 2.4 *Suppose that (X, d) is a finitely compact metric space whose isometries act transitively on X. Then every isometry of (X, d) is a motion.*

Proof Let f be any isometry of (X, d), and take any w in X. By transitivity, there is an isometry g such that $g\big(f(w)\big) = w$. Then, by Lemma 2.3, $gf(X) = X$ so that g is a bijection of X onto itself. Thus, $f(X) = g^{-1}(X) = X$, so f is a motion of X. ■

For example, the set of translations of a finite-dimensional normed vector space V acts transitively on V, so every isometry of V is a motion of V. As another illustration

https://doi.org/10.4153/S0008439520000867 Published online by Cambridge University Press

of Theorem 2.4, we refer to Busemann's text [9, pp. 345–346]. In the context of the G-spaces considered by Busemann (these are finitely compact metric spaces with modest, but natural, assumptions about distances between points; see [9, p. 37]), Busemann defines a *symmetry* of X about a point x_0 in X to be a motion f that is an involution (not the identity) of X onto itself such that x_0 is an isolated fixed point of f (intuitively, f is a "rotation" of order two about x_0). The space X is *symmetric* if such a symmetry exists at every point of X, and it is immediate for such spaces the group $\mathrm{M}(X)$ acts transitively on X (for given any points p and q, we can "rotate" the space about the midpoint of the segment $[p,q]$). Thus, all isometries of symmetric G-spaces are motions.

Theorem 2.4 gives a criterion for all isometries to be motions. By contrast, the next result gives a criterion that guarantees that a given isometry is a motion. In addition, this result frees us from the constraints of having a fixed point, and of mapping the space X into itself. We will use the notation $B_X(a,r)$ for the *closed* ball in X with centre a and radius r, and similarly for $B_Y(a,r)$.

Theorem 2.5 *Let (X,d) and (Y,d') be metric spaces, where (X,d) is finitely compact, and let $f\colon X\to Y$ be an isometry. Then $f(X)=Y$ if and only if there exists a point $a\in X$ such that for each $r>0$ there is an isometry ψ_r of $B_Y\big(f(a),r\big)$ into $B_X(a,r)$.*

Proof First, we suppose that the maps ψ_r exist and show that $f(X)=Y$. For brevity, for any positive r, let $B_X=B_X(a,r)$ and $B_Y=B_Y(f(a),r)$. We are given the existence of an isometry ψ_r of B_Y into B_X, and so, by Corollary 2.2, we see that, for any a in X, and any positive r, f maps B_X onto B_Y. Now take any y in Y and let $r=d_Y(f(a),y)$. Then $y\in B_Y=f(B_X)$, so that $f(X)=Y$. Finally, the reverse implication follows, for if $f(X)=Y$, we can choose any $a\in X$ and then take each ψ_r to be the restriction of the isometry f^{-1} to $B_Y(f(a),r)$. ∎

Observe that Theorem 2.5 (with $Y=X$) contains Lemma 2.3. Indeed, if f is an isometry of a finitely compact space (X,d) with a fixed point a, then, for each positive r, we can take each ψ_r in Theorem 2.5 to be the identity map so that $f(X)=X$. We can illustrate Theorem 2.5 in a concrete situation by examining the isometries and motions of the space $[0,+\infty)$ (with the Euclidean metric) into itself. Here, the maps $x\mapsto x+t$, where $t>0$, are isometries, but not motions.

We note that if Theorem 2.5 is applicable, then X and Y are homeomorphic (for then f is a homeomorphism), and this shows that the conclusion may fail if we only know that, for each a in X, an isometry $\psi_{a,r}\colon B_Y(f(a),r)\to B_X(a,r)$ exists for some positive r that depends on a. For example, we can take $X=\{0\}$, $Y=\{0,1\}$, and f the identity function on X. If Y is connected, then the suggested variant of Theorem 2.5 does hold.

Lemma 2.6 *Suppose that (X,d) and (Y,d') are metric spaces, where (X,d) is finitely compact and $f\colon X\to Y$ an isometry. Then $f(X)$ is closed in Y.*

Proof Select any y in the closure of $f(X)$. Then there is a sequence (x_n) in X with $f(x_n)\to y$. A convergent sequence is bounded, so there exists $R>0$ such that

$d'(f(x_1), f(x_n)) \leq R$ for all n. Then $d(x_1, x_n) = d'(f(x_1), f(x_n)) \leq R$ for all n. As X is finitely compact, the closed ball with centre x_1 and radius R is compact. Hence, (x_n) has a convergent subsequence. There is no harm in assuming $x_n \to x^* \in X$. Because f is continuous, $f(x_n) \to f(x^*)$. Hence, $y = f(x^*) \in f(X)$, so $f(X)$ is closed. ∎

Theorem 2.7 *Let (X, d) and (Y, d') be metric spaces, where (X, d) is finitely compact and (Y, d') is connected, and let $f: X \to Y$ be an isometry. Then the following are equivalent.*

(i) *f is a motion.*
(ii) *For each $a \in X$, there exists $r = r(a) > 0$ and an isometry $\psi_{a,r}$ of the closed ball $B_Y(f(a), r)$ into the closed ball $B_X(a, r)$.*
(iii) *f is an open mapping.*

Proof First, we show that (i) implies (ii). Because f is a motion, it is surjective, so f maps each closed ball $B_X(a, r)$ onto the closed ball $B_Y(f(a), r)$. Hence, $\psi_{a,r} = (f|B_X(a, r))^{-1}$ is an isometry of the required type.

Next, we demonstrate that (ii) implies (iii). We begin by verifying that f is an isometry of $B_X = B_X(a, r)$ onto $B_Y = B_Y(f(a), r)$. Corollary 2.2 gives that $f|B_X$ is a surjective isometry of B_X onto B_Y. In particular, $f|B_X$ maps each open ball in X with centre a and radius $\rho \in (0, r)$ onto the open ball in Y with centre $f(a)$ and radius ρ. Because this holds for all $a \in X$, the function f maps open sets in X onto open sets in Y.

Finally, (iii) entails (i). Because f is an open mapping, $f(X)$ is an open subset of Y. Lemma 2.6 implies that $f(X)$ is a closed subset of Y. Since Y is connected, $f(X) = Y$. ∎

3 Some Remarks

In this section, we make a few brief (and somewhat informal) remarks on similar ideas that interested readers may pursue at their leisure. First, we consider finitely compact metric spaces. Busemann [9, p. 6] defines a metric space to be (i) *compact* if every infinite subset has an accumulation point, and (ii) *finitely compact* if every bounded infinite subset has an accumulation point [9, p. 37], but a more modern approach is, of course, via open sets and sequential compactness. Here, we merely mention the standard results that a metric space is compact if and only if it is sequentially compact, and that it is finitely compact (in Busemann's sense) if and only if each closed ball is compact (which is our earlier definition). On [9, p. 403], Busemann remarks (with references) that a Hausdorff topological space can be metrised so as to become a finitely compact metric space if and only if it is locally compact and second countable, and it is known that a second countable topological space is compact if and only if it is sequentially compact [13, p. 235]. Finally, the interested reader can consult some of Busemann's ideas in [26].

Next, we consider the group of motions of a finitely compact metric space. If (X, d) is a finitely compact metric space, then the group $\mathrm{M}(X)$ of motions of X is a

https://doi.org/10.4153/S0008439520000867 Published online by Cambridge University Press

topological group. First, in [9, pp. 16–18] (see also [26, p. 118]) Busemann constructs a metric δ on the group $\mathrm{M}(X)$ of motions of a metric space (X, d) in such a way that the metric space $\big(\mathrm{M}(X), \delta\big)$ is finitely compact when (X, d) is finitely compact, and also compact when (X, d) is compact. For an alternative approach, see [27, p. 155], where $\mathrm{M}(X)$ is given the structure of a topological group via the compact-open topology on the space of continuous maps from X to itself.

The next topic is concerned with *non-expansive maps* and *local isometries*. Let f be a map of X into itself. Then f is (i) *non-expansive* if, for all x and y in X, $d(f(x), f(y)) \le d(x, y)$, and (ii) a *local isometry* if f is surjective, and if each x in X has an open neighbourhood on which f is an isometry. There are many interesting questions concerning these types of maps (for example, which conditions imply that a local isometry is a motion?), and we refer the reader to [9–11, 18–20] for more details.

Finally, suppose now that (X, d) is a finitely compact metric space whose isometries act transitively on X (so motions and isometries coincide). The *stabiliser* $\mathrm{S}(x)$ of a point x in X is the group $\{g \in \mathrm{M}(X)\colon g(x) = x\}$, and if $y = h(x)$, where h is an isometry, then $\mathrm{S}(y) = h\mathrm{S}(x)h^{-1}$; that is, $\mathrm{S}(x)$ and $\mathrm{S}(y)$ are conjugate subgroups. It follows that the entire group $\mathrm{M}(X)$ is determined by any set T of isometries that acts transitively on X, and the stabiliser $\mathrm{S}(\xi)$ of any given point ξ in X. Explicitly, if h is any isometry of X, then there is some t in T such that $t\big(h(\xi)\big) = \xi$, so that $h = t^{-1}g$ for some t in T and some g in $\mathrm{S}(\xi)$. In some important examples, T is a normal subgroup of $\mathrm{M}(X)$, and $\mathrm{T} \cap \mathrm{S}(\xi)$ contains only the identity map, and in this case, we say that $\mathrm{I}(X)$ is the *semi-direct product* of $\mathrm{S}(\xi)$ and T.

4 Some Examples

We now apply the ideas above to a variety of examples, in each of which an isometry is necessarily surjective. One of the most important tasks in an introduction to any of the three classical geometries (Euclidean, spherical, and hyperbolic) of constant curvature is to establish the group structure of the isometries, and, as we have seen, this depends on the surjectivity of the isometries. It is a standard exercise, and an elementary one, in linear algebra to show that every Euclidean isometry of $\mathbb{R}^n$ is the composition of a translation and an orthogonal (linear) map of $\mathbb{R}^n$ onto itself. However, even more is true, and in Section 5, we show that any isometry with respect to any norm on $\mathbb{R}^n$ is surjective.

Spherical geometry is the geometry of the unit sphere $\mathbb{S}^n$ in $\mathbb{R}^{n+1}$, and as $\mathbb{S}^n$ is compact, every isometry of $\mathbb{S}^n$ is surjective. Hyperbolic n-space can be taken to be the open unit ball $\mathbb{B}^n$ in $\mathbb{R}^n$, equipped with the hyperbolic distance h that is derived from the metric $2|dx|/(1 - \|x\|^2)$, and is given by

$$\sinh^2 \tfrac{1}{2}h(x, y) = \frac{\|x - y\|}{\sqrt{(1 - \|x\|^2)(1 - \|y\|^2)}}$$

[4, p. 40]. It is known that each closed hyperbolic ball is a closed Euclidean ball; thus, $(\mathbb{B}^n, h)$ is finitely compact. Further, as the inversion across any Euclidean hypersphere that is orthogonal to the unit sphere $\mathbb{S}^{n-1}$ is a hyperbolic isometry, we see that the hyperbolic isometries act transitively on $\mathbb{B}^n$. Thus, every hyperbolic

https://doi.org/10.4153/S0008439520000867 Published online by Cambridge University Press

isometry is surjective. The desired result can be achieved (in all three geometries and in all dimensions) by, for example, applying Theorem 2.4, having first established the existence of reflections across geodesics, planes, *etc*. as appropriate. Indeed, it is also advantageous to show that every motion is a composition of such reflections. An alternative approach (which has been mentioned above and also applies to all three geometries) is to show that each geometry supports an isometry that is a "rotation of order two" about any point of the space.

We discuss one more example in this section, namely, the *quasi-hyperbolic plane*, which appears in a number of different situations. We begin with the complex plane $\mathbb{C}$ equipped with the Euclidean metric, and the cyclic group T of translations generated by $z \mapsto z + 2\pi i$. From a topological perspective, the quotient space $\mathbb{C}/T$ is a Euclidean cylinder (obtained by identifying the sides $y = 0$ and $y = 2\pi$ of the strip $\{x + iy: 0 \leq y \leq 2\pi\}$), but from a complex analytic perspective, $\mathbb{C}/T$ is a Riemann surface that is conformally equivalent to the punctured plane $\mathbb{C}\backslash\{0\}$ (which is the image of its universal covering space $\mathbb{C}$ by the universal covering map $z \mapsto \exp z$). We will consider $\mathbb{C}/T$ from both points of view.

From the perspective of complex analysis, the *quasi-hyperbolic plane* is the set $\mathbb{C}\backslash\{0\}$ equipped with the quasi-hyperbolic distance μ derived from the *quasi-hyperbolic* metric $|dz|/|z|$ (which is the local projection of the Euclidean metric on its universal covering space $\mathbb{C}$). As any bounded set in $(\mathbb{C}\backslash\{0\}, \mu)$ lies in some annulus $\{z: r \leq |z| \leq R\}$, where $0 < r \leq R < +\infty$, this metric space is finitely compact. Further, as Euclidean rotations about the origin, and maps $z \mapsto kz$, where $k \neq 0$, are μ-isometries, the isometries act transitively on $\mathbb{C}\backslash\{0\}$, so each isometry is surjective. Finally, the reader should note that this is not the same as the quasi-hyperbolic plane that is defined in [9].

Let us now identify the quotient space $\mathbb{C}/T$ with the vertical Euclidean cylinder $\mathcal{C} = \mathbb{S}^1 \times \mathbb{R}$ in $\mathbb{R}^3$, and give $\mathcal{C}$ the distance γ obtained by minimizing the Euclidean arc length of paths on $\mathcal{C}$. As vertical translations, and rotations about the vertical axis, are isometries of $\mathcal{C}$, the set of isometries acts transitively, so every isometry of $\mathcal{C}$ is surjective. In fact, it can be shown that the quasi-hyperbolic length of a path in $\mathbb{C}\backslash\{0\}$ is the same as the Euclidean length of its image on $\mathcal{C}$.

The Euclidean isometries of the vertical cyclinder $\mathcal{C}$ arise in another context. A *frieze group* is a discrete group F of isometries of $\mathbb{C}$ that leaves the real axis $\mathbb{R}$ invariant, and has the property that its subgroup T of translations is an infinite cyclic group (see, for example, [1, 23, 24]). As two frieze groups are identified if they are conjugate by an affine map, we can assume that T is generated by $z \mapsto z + 2\pi$. Now, T is a normal subgroup of F, and by analyzing the structure of frieze groups, it follows that each quotient group F/T is a subgroup of the Klein four-group, and from this (see [5]), we can easily derive the familiar result that, up to conjugation, there are only seven possible frieze groups. This algebraic proof is much shorter than the usual arduous geometric proof. Now we can identify the quotient space $\mathbb{C}/T$ with the cylinder $\mathcal{C}$ and the quotient group F/T with a group of Euclidean isometries of $\mathcal{C}$. As F acts discretely on $\mathbb{C}$, and the elements of F/T leave the circle $\mathbb{S}^1 \times \{0\}$ (which corresponds to the real axis in $\mathbb{C}$) invariant, we find that F/T is a *finite* group of isometries of $\mathcal{C}$; thus, frieze groups can be identified with certain finite subgroups of the isometry group of the quasi-hyperbolic plane.

Finally, we note that, up to a Euclidean isometry, the cylinder, the unit sphere, and Euclidean planes are the only surfaces in $\mathbb{R}^3$ that support a transitive group of Euclidean isometries [15, p. 218]. More generally, on [9, p. 371], Busemann gives a complete list of G-surfaces that possess a transitive group of motions (but a discussion of this would take us too far afield).

5 Minkowski Spaces

A finite-dimensional, real, normed vector space is called a *Minkowski space*; see [26, p. 146]. The Mazur–Ulam theorem states that if E and F are real, normed vector spaces, but not necessarily of finite dimension, and f is a surjective isometry of E onto F, then f is an affine map; that is, the map $x \mapsto f(x) - f(0)$ is a linear transformation. An elegant, short, and very elementary proof of this in [25] shows that, under the given hypothesis, if $a, b \in E$, and $t \in [0,1]$, then $f((1-t)a + tb) = (1-t)f(a) + tf(b)$. This, in turn, implies that f is affine. The restriction to surjective isometries here is necessary; for example, if $\mathbb{R}^2$ has norm $(x, y) \mapsto \max\{|x|, |y|\}$, then the map $f{:}\,\mathbb{R} \to \mathbb{R}^2$ defined by $f(x) = (x, \sin x)$ is an isometry that is not affine [29]. However, a closed ball in a finite-dimensional normed space is compact [7, p. 63], and the set of translations of a normed space acts transitively on that space. Thus, Theorem 2.4 and the Mazur–Ulam theorem show that every isometry of a Minkowski space E into itself is surjective, and therefore affine. For a general, elementary, account of these ideas, see [21].

These remarks lead naturally to a discussion of different norms on Euclidean space $\mathbb{R}^n$, and while it is an elementary fact from linear algebra that every isometry of $\mathbb{R}^n$ with the Euclidean norm is surjective and affine, the argument given above shows that an isometry of $\mathbb{R}^n$ with respect to any norm is surjective and affine. In this discussion, a point x in $\mathbb{R}^n$ is a row vector; its transpose x^{t} is a column vector, and the Euclidean norm, distance, and inner product are denoted by $\|x\|$, $\|x - y\|$, and $x \cdot y$, respectively. It is well known that if $1 \le p \le +\infty$, $x = (x_1, \dots, x_n)$, and

$$\|x\|_p = \Big(\sum_{j=1}^{n} |x_j|^p\Big)^{1/p},$$
$$\|x\|_\infty = \max\{|x_j| : 1 \le j \le n\},$$

then $\|x\|_p$ is a norm on $\mathbb{R}^n$ with $d_p(x, y) = \|x - y\|_p$ the associated distance function. Of course, for $p = 2$, this is the usual Euclidean norm and distance. The case $p = 1$ is also popular, for in this case, the metric on $\mathbb{R}^n$ is the so-called *taxicab distance* d_1. The elementary nature of this case has led to many instances of publications devoted to exploring the isometries of these distances, sometimes of d_1 for small values of n, sometimes of d_1 for all n, and sometimes for d_p; see, for example, [8, 12, 14, 17, 22, 28], and doubtless, this list could be longer.

Simple descriptions of the isometries of each of the norms $\|x\|_p$ are known. As the case $p = 2$ leads to the familiar group of orthogonal matrices, let us assume that $p \neq 2$. We now give the isometry groups of the norms $\|x\|_p$ when $p \neq 2$. A real $n \times n$ matrix is a *permutation matrix* if and only if it has exactly one non-zero entry in each row and each column, and each such element is 1; a real $n \times n$ matrix is a *signed permutation matrix* if and only if it has exactly one non-zero entry in each row and each column, and each such element is either 1 or −1. Because a determinant changes by a factor −1 when we interchange two adjacent rows, we see that each such matrix (of either type) is non-singular. Both types of matrices form a multiplicative group, and a simple combinatorial argument shows that these groups have order $n!$ and $2^n n!$, respectively. In fact, the signed permutation matrices are precisely the matrices in the (Euclidean) orthogonal group that have integer entries. Indeed, it is easy to see that a signed permutation matrix is orthogonal. Conversely, if A is an orthogonal matrix, then each row and column has (as a vector) length 1, so if the entries of A are integers, then each row and column can only have one non-zero entry, and this entry must be ±1. Finally, if A is a signed permutation matrix, $x \in \mathbb{R}^n$, and $y^{\mathrm{t}} = Ax^{\mathrm{t}}$, then $(y_1, \dots, y_n) = \big(\varepsilon_1 x_{\rho(1)}, \dots, \varepsilon_n x_{\rho(n)}\big)$ for some permutation ρ of $\{1, \dots, n\}$, and for some choice of each ε_j from $\{-1, 1\}$. It is clear from this that any signed permutation matrix is a linear isometry of $\big(\mathbb{R}^n, d_p\big)$. In fact, for $p \neq 2$, these are the only linear isometries of this space. This can be proved by elementary arguments, but we will present a more revealing proof (see Theorem 5.3) after discussing the John inellipsoid and John circumellipsoid for a convex body relative to an arbitrary norm on $\mathbb{R}^n$.

An elementary example shows that the isometries of a norm on $\mathbb{R}^n$ need not be Euclidean isometries. For example, consider the norm $\|(x, y)\|^* = 2|x| + |y|$ on $\mathbb{R}^2$. The linear map $f\colon \mathbb{R}^2 \to \mathbb{R}^2$ given by $(x, y) \mapsto (-y/2, 2x)$ preserves the norm $\|\cdot\|^*$, and so provides an isometry of $\mathbb{R}^2$ with this norm that is not a Euclidean isometry.

Let us now consider a general norm $\|x\|^*$ on $\mathbb{R}^n$. Then the closed unit ball $\overline{B}^* = \{x \in \mathbb{R}^n \colon \|x\|^* \leq 1\}$ for this norm is a compact convex subset of $\mathbb{R}^n$ that is symmetric about about the origin (that is, it is invariant under the antipodal map $x \mapsto -x$), and, because of this, both the identity map and the antipodal map are isometries. In fact, any compact convex set in $\mathbb{R}^n$ that has the origin as an interior point and that is invariant under the antipodal map is the closed unit ball for some norm on $\mathbb{R}^n$; this follows from a construction due to Minkowski; see [26, p. 156]. In particular, a closed regular n-gon in $\mathbb{R}^2$ with centre at the origin is the closed unit ball for a norm on $\mathbb{R}^2$ if and only if n is even.

We will now introduce the concept of *John ellipsoids*. An *ellipsoid* in $\mathbb{R}^n$ is the image of the closed Euclidean unit ball $\overline{\mathbb{B}}^n = \{x \in \mathbb{R}^n \colon \|x\| \leq 1\}$ (with respect to the Euclidean norm $\|x\|$) under a bijective linear map of $\mathbb{R}^n$ onto itself. A *convex body* is a compact convex set with nonempty interior. In 1948, F. John proved that each convex body K in $\mathbb{R}^n$ contains a unique ellipsoid of maximal volume, and, likewise, there is a unique ellipsoid with minimal volume, which contains K [16]. These ellipsoids are now called the *John inellipsoid*, and the *John circumellipsoid* for K. When $n = 2$, the terms *Steiner inellipse* and *Steiner circumellipse* are customary.

Lemma 5.1 *Suppose that $\|x\|^*$ is a norm on $\mathbb{R}^n$ with closed unit ball K for $\|x\|^*$. If either the John circumellipsoid or the John inellipsoid for K is the closed unit ball $\overline{\mathbb{B}}^n$, then every isometry of $\|x\|^*$ is a Euclidean isometry.*

https://doi.org/10.4153/S0008439520000867 Published online by Cambridge University Press

Proof It is not difficult to prove that if f is a bijective linear map of $\mathbb{R}^n$ onto itself, K is a convex body and E is its John circumellipsoid (or John inellipsoid), then $f(K)$ is a convex body and $f(E)$ is its John circumellipsoid (or John inellipsoid). The Mazur–Ulam theorem implies that every isometry for $\|x\|^*$ is an affine mapping. First, suppose that $\overline{\mathbb{B}}^n$ is the circumellipsoid for K and f is an isometry of $\|x\|^*$. Then f is affine and $f(\overline{\mathbb{B}}^n) = \overline{\mathbb{B}}^n$. The only affine self-maps of $\overline{\mathbb{B}}^n$ are the Euclidean isometries, so f is a Euclidean isometry. The argument when $\overline{\mathbb{B}}^n$ is the John inellipsoid is analogous. ∎

To complete our discussion, we need the following elegant characterization of those cases in which the closed Euclidean unit ball $\overline{\mathbb{B}}^n$ is the John inellipsoid or John circumellipsoid of a convex body. For a proof of this result, see especially [3, p. 13], but also [2, 7, 16], and for a general perspective on these ideas, see [6].

Theorem 5.2 *Each convex body K contains a unique ellipsoid of maximal volume. This ellipsoid is $\overline{\mathbb{B}}^n$ if and only if $\overline{\mathbb{B}}^n \subset K$ and, for some integer m, there are points $u_1, \dots, u_m$ in $\mathbb{S}^{n-1} \cap \partial K$, where $\mathbb{S}^{n-1} = \partial\overline{\mathbb{B}}^n$, and positive numbers $c_1, \dots, c_m$, such that*

$$c_1 u_1 + \cdots + c_m u_m = 0, \tag{5.1}$$

and, for all x in $\mathbb{R}^n$,

$$x = \sum_{j=1}^{m} c_j \big(x \cdot u_j\big) u_j. \tag{5.2}$$

Likewise, K is contained in a unique ellipsoid of minimal volume that is $\overline{\mathbb{B}}^n$ if and only if $K \subset \overline{\mathbb{B}}^n$, and, for some integer m, there are points $u_1, \dots, u_m$ in $\mathbb{S}^{n-1} \cap \partial K$, and positive numbers $c_1, \dots, c_m$, such that (5.1) and (5.2) hold.

We now determine the isometry group for the norm $\|\cdot\|_p$ when $1 \le p \le +\infty$ and $p \ne 2$ by making use of Theorem 5.2.

Theorem 5.3 *If $1 \le p \le +\infty$ and $p \ne 2$, then the isometry group $I(\mathbb{R}^n, d_p)$ is the group of $n \times n$ signed permutation matrices.*

Thus, for $p \ne 2$ the group of isometries for d_p is independent of p, and $f\colon \mathbb{R}^n \to \mathbb{R}^n$ is an isometry of $\|\cdot\|_p$ if and only if there is a signed permutation matrix A with $f(x) = Ax^t$, where $x = (x_1, \dots, x_n)$.

Proof Immediately after introducing signed permutation matrices, we observed that for any signed permutation matrix A, the map $x \mapsto Ax^t$ belongs to $\mathrm{I}(\mathbb{R}^n, d_p)$. It remains to show that every isometry of d_p is given by a permutation matrix. Let $\overline{B}_p$ denote the closed unit ball for $\|\cdot\|_p$ on $\mathbb{R}^n$; of course, $\overline{B}_2 = \overline{\mathbb{B}}^n$.

We will use Theorem 5.2 with $K = \overline{B}_p$, $m = 2n$,

$$c_1 = \cdots = c_{2n} = \tfrac{1}{2}, \quad \text{and} \quad \{u_1, \dots, u_{2n}\} = \{e_1, -e_1, \dots, e_n, -e_n\}.$$

Because $\|u_j\|_p = 1 = \|u_j\|_2$, it is evident that $u_j \in \mathbb{S}^{n-1} \cap \partial \overline{B}_p$ for all j. As $(1/2)e_j + (1/2)(-e_j) = 0$ for all j, it is clear that (5.1) holds. For $x \in \mathbb{R}^n$, the identity $x = \sum_{j=1}^n (x \cdot e_j) e_j$ implies that (5.2) holds with $c_j = 1/2$ for all j.

The remainder of the proof naturally separates into the cases $2 < p \le +\infty$ and $1 \le p < 2$. In both cases, we use the elementary inequality $\|x\|_q \le \|x\|_p$, which holds for $0 < p < q \le +\infty$. First, suppose that $2 < p \le +\infty$. The inequality $\|x\|_p \le \|x\|_2$ implies that $\overline{\mathbb{B}}^n \subset \overline{B}_p$. Therefore, Theorem 5.2 implies that $\overline{\mathbb{B}}^n$ is the John inellipsoid for the convex body $\overline{B}_p$. Then Lemma 5.1 guarantees that every d_p-isometry is a Euclidean isometry.

Next, we prove that every d_p-isometry is given by a signed permutation matrix. We show that

$$\{e_1, -e_1, \ldots, e_n, -e_n\} = \partial \overline{B}_p \cap \mathbb{S}^{n-1}. \tag{5.3}$$

We already know that the left side is a subset of the right side. In order to prove (5.3), it is sufficient to show that if $\|x\|_2 = 1 = \|x\|_p$, equivalently, if

$$x_1^2 + \cdots + x_n^2 = 1 = |x_1|^p + \cdots + |x_n|^p, \tag{5.4}$$

then $x = \pm e_j$ for some j. First, suppose that $2 < p < +\infty$. If two of the x_j are non-zero, we can assume that they are x_1 and x_2. Then $|x_1|^p < |x_1|^2 < 1$ and $|x_2|^p < |x_2|^2 < 1$ while $|x_j|^p \le |x_j|^2$ for $j = 3, \ldots, n$. This violates (5.4), so if $2 < p < +\infty$, then $x = \pm e_j$ for some j. We leave the case $p = \infty$ to the reader. The fact that the identity (5.3) holds implies that any Euclidean isometry of d_p leaves the set $\{e_1, -e_1, \ldots, e_n, -e_n\}$ invariant, and it is easy to see that this implies that f is a signed permutation matrix.

Now assume that $1 \le p < 2$. Then $\|x\|_2 \le \|x\|_p$ and so $\overline{B}_p \subset \overline{\mathbb{B}}^n$. Hence, Theorem 5.2 implies that $\overline{\mathbb{B}}^n$ is the John circumellipsoid for B_p, and Lemma 5.1 guarantees that every d_p-isometry is a Euclidean isometry. As in the case $2 < p$, it suffices to show that (5.3) holds, and for this we need to verify that if (5.4) holds, then $x = \pm e_j$. If two of the x_j are non-zero, we can assume that they are x_1 and x_2. Then $|x_1|^2 < |x_1|^p < 1$ and $|x_2|^2 < |x_2|^p < 1$ while $|x_j|^2 \le |x_j|^p$ for $j = 3, \ldots, n$. This contradicts (5.4), so if $1 \le p < 2$, then $x = \pm e_j$ for some j. As in the case $2 < p$, this implies that when $1 \le p < 2$, every isometry of d_p is a signed permutation matrix. ∎

6 Regular Polygons

In a similar elementary manner, we determine in the isometry group for a norm on $\mathbb{R}^2$ with closed unit ball equal to the closed regular $2k$-gon $\mathcal{P}_{2k}$ inscribed in the unit circle and having one vertex at $(1, 0)$. It is simpler to use complex notation in this situation, so we do.

Let $\|z\|^*$ be a norm on $\mathbb{C}$, the complex plane, with closed unit ball $\mathcal{P}_{2k}$. Geometrically, it is evident that $\overline{\mathbb{B}}^2$ is the John circumellipse for $\mathcal{P}_{2k}$, so Lemma 5.1 implies that every isometry for $\|z\|^*$ is given by an orthogonal matrix. We explicitly verify that $\overline{\mathbb{B}}^2$ is the John circumellipse. The set $\mathbb{S}^1 \cap \partial \mathcal{P}_{2k}$ consists of the vertices of $\mathcal{P}_{2k}$; that is, $u_j = \exp(j2\pi i/(2k)) = \exp(j\pi i/k) = \omega^j$, $0 \le j \le 2k-1$, where $\omega = \exp(\pi i/k)$ We explicitly verify (5.1) and (5.2). Because the centre of mass of the $(2k)$-th roots of unity

https://doi.org/10.4153/S0008439520000867 Published online by Cambridge University Press

is the origin, (5.1) holds for any choice of the same positive constant c_j for all j. Now we verify (5.2). For $z \in \mathbb{C}$,

$$\sum_{j=0}^{2k-1} (z \cdot u_j) u_j = \sum_{j=0}^{2k-1} (z \cdot \omega^j) \omega^j = \sum_{j=0}^{2k-1} \operatorname{Re}\left(z\overline{\omega}^j\right) \omega^j$$
$$= \frac{1}{2} \sum_{j=0}^{2k-1} \left(z\overline{\omega}^j + \overline{z}\omega^j\right) \omega^j = kz.$$

Thus, $c_j = 1/k$, $0 \le j \le 2k-1$, works in (5.2).

Theorem 6.1 *The group $I(\mathbb{C}, d_*)$ is the dihedral group D_{2k}, the Euclidean group of isometries for $\mathcal{P}_{2k}$.*

Proof Because the closed unit disc is the John circumellipse for $\mathcal{P}_{2k}$, every isometry for d_* is given by an orthogonal matrix that leaves the set $\overline{\mathbb{B}}^2$ invariant. This Euclidean isometry leaves the polygon $\mathcal{P}_{2k}$ invariant, so it belongs to the dihedral group D_{2k}, which is the Euclidean symmetry group of $\mathcal{P}_{2k}$ and has $4k$ elements. Conversely, we show that each $f \in D_{2k}$ is an isometry of d_*. Note that $f \in D_{2k}$ maps $\partial\mathcal{P}_{2k}$ onto itself, so $\|u\|^* = 1$ and $\|f(u)\|^* = 1$ for all $u \in \partial\mathcal{P}_{2k}$. For any nonzero $z \in \mathbb{C}$, $u = z/\|z\|^*$ lies on $\partial\mathcal{P}_{2k}$. Therefore, for nonzero $z \in \mathbb{C}$, $z = tu$, where $t = \|z\|^* \ge 0$, and so

$$\|f(z)\|^* = \|f(tu)\|^* = \|tf(u)\|^* = t\|f(u)\|^* = t = \|tu\|^* = \|z\|^*.$$

Hence, f is also a d_* isometry. ∎

When k is even, this result is a special case of the classification of symmetric norms on $\mathbb{R}^n$ [21, Thm. 1]. A norm $\|x\|^*$ on $\mathbb{R}^n$ is *symmetric* if $\|Px\|^* = \|x\|^*$ for all $x \in \mathbb{R}^n$ and all $n \times n$ signed permutation matrices P. When $n = 2$, a symmetric norm satisfies $\|\pm e_1\|^* = \|\pm e_2\|^*$. If $\|x\|^*$ has closed unit ball equal to $\mathcal{P}_{2k}$, then $\pm e_2$ are vertices for $\mathcal{P}_{2k}$ if and only if k is even. Consequently, [21, Thm. 1] only gives the isometry groups for symmetric norms with closed unit ball is equal to $\mathcal{P}_{4k}$. Our elementary approach produces the answer for any norm with unit ball $\mathcal{P}_{2k}$, $k \ge 1$.

Acknowledgment The authors would like to thank the referee for carefully reading the original version of this paper and making a number of useful suggestions that improved the paper.

References

[1] M. A. Armstrong, *Groups and symmetry*. Undergraduate Texts in Mathematics, Springer-Verlag, Berlin, 1988. https://doi.org/10.1007/978-1-4757-4034-9

[2] K. Ball, *Ellipsoids of maximal volume in convex bodies.* Geom. Dedicata **41**(1992), 241–250. https://doi.org/10.1007/bf00182424

[3] K. Ball, *An elementary introduction to modern convex geometry.* In: Flavours of geometry, Math. Sci. Res. Inst. Publ., 31, Cambridge University Press, Cambridge, UK, 1997, pp. 1–58. https://doi.org/10.2977/prims/1195164788

[4] A. F. Beardon, *The geometry of discrete groups*. Springer-Verlag, New York, 1983. https://doi.org/10.1007/978-1-4612-1146-4

[5] A. F. Beardon, *Frieze groups, cylinders and quotient groups.* Math. Gazette **97**(2013), 95–100. https://doi.org/10.1017/s0025557200005465

[6] M. Berger, *Convexity*. Amer. Math. Monthly **97**(1990), 650–678. https://doi.org/10.1007/978-3-540-70997-8

[7] B. Bollobás, *Linear analysis*. 2nd ed., Cambridge Univ. Press, Cambridge, UK, 2012. https://doi.org/10.1017/cbo9781139168472

[8] V. G. Bulgarean, *Study of isometry groups*. Ph.D. dissertation, Babes-Bolyai University, Cluj-Napoca, 2014.

[9] H. Busemann, *The geometry of geodesics*. Dover (reprint), New York, NY, 2005.

[10] A. Całka, *Local isometries of compact metric spaces.* Proc. Amer. Math. Soc. **85**(1982), 643–647. https://doi.org/10.2307/2044083

[11] A. Całka, *On conditions under which isometries have bounded orbits.* Colloq. Math. **48**(1984), 219–227. https://doi.org/10.4064/cm-48-2-219-227

[12] T. Ermiş and R. Kaya, *On the isometries of 3-dimensional maximum space*. Konuralp J. Math. 3(2015), 103–114.

[13] P. Fitzpatrick and H. Royden, *Real analysis*. 4th ed., Prentice-Hall, Boston, MA, 2010.

[14] Ö. Gelişgen and R. Kaya, *The taxicab space group.* Acta Math. Hungar. **122**(2009), 187–200. https://doi.org/10.1007/s10474-008-8006-9

[15] H. W. Guggenheimer, *Differential geometry*. Dover Books on Advanced Mathematics, Dover Publications, New York, NY, 1977.

[16] F. John, *Extremum problems with inequalities as subsidiary conditions.* In: Studies and essays presented to R. Courant on his 60th birthday, January 8, 1948, Interscience Publishers Inc., New York, NY, 1948, pp. 187–204. https://doi.org/10.1007/978-1-4612-5412-6

[17] R. Kaya, Ö. Gelişgen, S. Ekmerçi, and A. Bayar, *On the group of isometries of the plane with generalized absolute value metric.* Rocky Mountain J. Math. **39**(2009), 591–603. https://doi.org/10.1216/rmj-2009-39-2-591

[18] W. A. Kirk, *On locally isometric mappings of a G-space on itself*. Proc. Amer. Math. Soc. **15**(1964), 584–586. https://doi.org/10.2307/2034752

[19] W. A. Kirk, *On conditions under which local isometries are motions.* Colloq. Math. **22**(1971), 229–232. https://doi.org/10.4064/cm-22-2-229-232

[20] W. A. Kirk, *A theorem on local isometries.* Proc. Amer. Math. Soc. **17**(1964), 453–455. https://doi.org/10.1090/S0002-9939-1966-0190886-0

[21] C.-K. Li, *Norms, isometries and isometry groups.* Amer. Math. Monthly **107**(2000), 334–340. https://doi.org/10.1080/00029890.2000.12005201

[22] C.-K. Li and W. So, *Isometries of lp-norm.* Amer. Math. Monthly **101**(1994), 452–453. https://doi.org/10.1080/00029890.1994.11996972

[23] R. C. Lyndon, *Groups and geometry*. London Mathematical Society Lecture Notes Series, 101, Cambridge University Press, Cambridge, UK, 1985. https://doi.org/10.1017/cbo9781107325685

[24] G. E. Martin, *Transformation geometry. An introduction to symmetry*. Undergraduate Texts in Mathematics, Springer-Verlag, New York-Berlin, 1982. https://doi.org/10.2307/3616871

[25] B. Nica, *The Mazur–Ulam theorem.* Exp. Math. **30**(2012), 397–398. https://doi.org/10.1016/j.exmath.2012.08.010

[26] A. Papadopoulos, *Metric spaces, convexity and nonpositive curvature*. 2nd ed., IRMA Lectures in Mathematics and Theoretical Physics, 6, European Math. Soc., Zürich, 2014. https://doi.org/10.4171/010

[27] J. G. Ratcliffe, *Foundations of hyperbolic manifolds*. Graduate Texts in Mathematics, 149, Springer-Verlag, New York, NY, 1994. https://doi.org/10.1007/978-3-030-31597-9_10

[28] D. Schattschneider, *The taxicab group.* Amer. Math. Monthly **91**(1984), 423–428. https://doi.org/10.1080/00029890.1984.11971453

[29] J. Väisälä, *A proof of the Mazur–Ulam theorem.* Amer. Math. Monthly **110**(2003), 633–635. https://doi.org/10.2307/3647749

Centre for Mathematical Sciences, University of Cambridge, Wilberforce Road, Cambridge CB3 0WB, UK
e-mail: afb@dpmms.cam.ac.uk

Department of Mathematical Sciences, University of Cincinnati, Cincinnati, OH 45221-0025, USA
e-mail: minda@ucmail.uc.edu