Distances and Homeomorphisms

Problem. Characterize continuous *alif* distances $d : U \times U \to \mathbb{R}_+$ on a convex bounded open subset U of $\mathbb{R}^n$, i.e., those for which the following is true:

($\aleph$) *Let $f : U \to U$ be any homeomorphism of U which extends by the identity map of its boundary ∂U. Then, if f is Lipschitz with respect to the distance d, it is Lipschitz with respect to the Euclidean distance e of U.*

Here Lipschitz is short for 'uniformly bi-Lipschitz', i.e., the distortion of distance by the homeomorphism as well as its inverse is bounded globally, i.e., there is a constant $L < \infty$ such that $d(f(P), f(Q)) \leq Ld(P, Q)$ and $d(f^{-1}(P), f^{-1}(Q)) \leq Ld(P, Q)$ hold for all pair of points P, Q of U.

We'll also look at some variants of this problem. For example, if the distance d is such that the Euclidean diameter of d-balls of a bounded radius is arbitrarily small near ∂U, then a homeomorphism $f : U \to U$ extends by the identity map of ∂U if it is *bounded* with respect to d, i.e., there is a constant $\delta < \infty$ such that $d(f(P), P) \leq \delta$ for any point $P \in U$; but, bounded homeomorphisms merit attention even when they don't extend by the identity of ∂U.

Convexity of U, i.e., closure with respect to shortest Euclidean path between pairs of points in it, reduces us to checking that length of small germs of segments or *elements* is distorted by at most the factor L :- For, using compactness, we can subdivide any segment PQ into finitely many such small segments $P_{i-1}P_i$ where $P_0 = P$ and $P_t = Q$. So each segment $f(P_{i-1})f(P_i)$ is at most L times longer than $P_{i-1}P_i$. Hence the broken line from $f(P) = f(P_0)$ to $f(Q) = f(P_t)$ formed by these t segments, and so à fortiori the segment $f(P)f(Q)$ of U, is at most L times longer than PQ. $\square$

We'll often denote the Euclidean length $e(P, Q)$ of a segment PQ also by PQ. So the distance axioms for e read: $PP = 0$, $PQ > 0$ if $P \neq Q$. $PQ = QP$, $PQ + QR \geq PR$. Likewise $d(P, P) = 0$, $d(P, Q) > 0$ if $P \neq Q$. $d(P, Q) = d(Q, P)$, $d(P, Q) + d(Q, R) \geq d(P, R)$; besides d is continuous which is necessary and sufficient to ensure that it gives the same topology on U:-

Continuity of d shows any open d-ball of U is e-open. Let C be any closed e-ball of $\mathbb{R}^n$ contained in U. It is e-compact. So C is also compact in the coarser metric, so Hausdorff, d-topology. So the two topologies coincide on C. In particular $\text{int}(C)$, an open e-ball, is d-open. The result follows because any e-open subset of U is a union of such open e-balls. $\square$

Our problem hinges on the conversion or comparison ratios $\frac{d(P,Q)}{PQ}$ between these two ways of measuring separation between pairs of distinct nearby points. As long as these positive numbers stay well away from 0 and ∞ intuition tells us all should be hunky-dory, but in case this is not so, the answer will involve on how these positive numbers approach 0 or ∞.

The one-dimensional case $n = 1$ of our problem is already very interesting and reveals almost all its features. So we'll linger on, and around, this case for a while to firm up our intuition into something more exact.

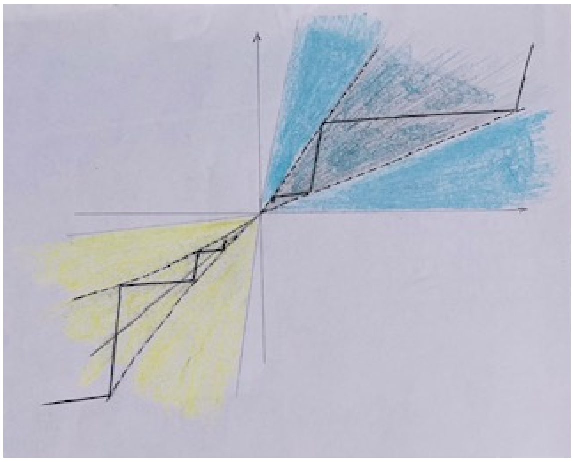
Now $U \subseteq \mathbb{R}$ is an open interval (a, b) where possibly $a = -\infty$ and/or $b = \infty$. One might think that continuity would sharpen the triangle inequality of d to additivity $d(P, Q) + d(Q, R) = d(P.R) \forall P < Q < R$ in this case, but this is seldom true. Indeed for $m > 1$ the triangle inequality is generally strict for the following class of natural distances on the interval:-

Any *parametrised arc of* $\mathbb{R}^m$, i.e., $1-1$ continuous function $d : (a, b) \to \mathbb{R}^m$, defines a continuous distance $d(P, Q) = d(P)d(Q)$ on the interval.

However for $m = 1$ a distance defined thus is additive. It remains unchanged if we add a constant or multiply this function by -1, so we can assume $d : U \to \mathbb{R}$ strictly increasing and zero at any chosen base point $O \in U$. So $d(O, P) = |d(P)|$ and $d(P, Q) = d(Q) - d(P)$ if $P < Q$. The conversion ratios become the slopes $\frac{d(Q)-d(P)}{Q-P}$ of the chords of the graph of this function. □

An orientation preserving homeomorphism $f : U \to U$ is also the same as a strictly increasing continuous function, but with the additional requirement of surjectivity $f(U) = U$. Indeed the inverse of any strictly monotone continuous function is automatically continuous. A strictly decreasing continuous surjective function, i.e., an orientation reversing homeomorphism of the interval U, has exactly one *fixed point*, that is, its graph cuts the diagonal of $U \times U$ exactly once. On the other hand, a strictly increasing homeomorphism can have any closed subset $F \subseteq U$ as its fixed-point-set $\{P \in U : f(P) = P\}$. The open complement of F is the disjoint union of at most countably many intervals (a_k, b_k) on each of which either always $f(P) < P$ or $f(P) > P$. □

Basically because any $O \in (a, b)$ can be an a_k or b_k of some f, an $\aleph$-distance d has to be more than just continuous at all points. To get at what the optimum condition might be let us consider the *limits of slope* $\frac{d(Q)-d(P)}{Q-P}$ of chords of the continuous strictly increasing zigzag graph below:-



At each corner between a *zig,* a segment with a small positive slope m, and a *zag,* a segment with a big slope M, this limit is m or M if P and Q approach the corner from that side. But if no constraint is put any value in $[m, M]$ can occur as a limit. So, at that exceptional point O, it is exactly all values in the interval $[\liminf_r(m), \limsup_r(M)]$ that occur as limits if we constrain (P,Q) to the right of O. All rays with these slopes constitute that blue *forward cone,* which can well be the entire closed first quadrant. Likewise, the yellow *backward cone* indicates all limits at O of $\frac{d(Q)-d(P)}{Q-P}$ under the constraint $P < Q \leq O$. The smallest interval containing these one-sided limits, i.e., $[\liminf(m), \limsup(M)]$, gives all unconstrained limits of slope, and can be visualized by a *two-sided cone* (not drawn) of all lines through O with these slopes.

The darker forward subcone shows all *right derivatives* of d at O, i.e., limits of slope under the constraint $O = P < Q$ which is strictly tighter than '(P,Q) to the right of O'. The biggest and smallest of these, the *right Dini derivatives* of d, are the slopes of the two lines on which all zigzag corners to the right of O lie. On the other hand to the left of O the zigzag corners lie on two curves making a cusp at O, so d is *left differentiable,* its sole left derivative at O being the slope of the common tangent to the two curves at this cusp. Finally, the smallest interval containing all right and left derivatives at O comprises all *neutral derivatives* of d at O, i.e., limits of slope taken under the constraint $O \in [P,Q]$, and can be visualized as a two-sided closed sub-cone. □

These definitions apply more generally, e.g., for any function $d : (a,b) \to \mathbb{R}$, if $\frac{d(Q)-d(P)}{Q-P}, O \in [P,Q]$ is arbitrarily close to ℓ infinitely often in a sufficiently small neighbourhood of O, then ℓ is a neutral *derivative* of d at O, and if there is a unique such ℓ then d is *differentiable* at O:- This is equivalent to the usual definition of differentiability of d at O because $\frac{d(Q)-d(P)}{Q-P}, O \in (P,Q)$ is a convex linear combination of $\frac{d(O)-d(P)}{O-P}$ and $\frac{d(Q)-d(O)}{Q-O}$. □

Examples. Consider any doubly infinite increasing sequence $P_i \in (a,b), i \in \mathbb{Z}$ with no limit point in (a,b). The fixed point free homeomorphism f which maps each segment P_iP_{i+1} linearly on $P_{i+1}P_{i+2}$ is (uniformly bi-)Lipschitz with respect to the Euclidean distance if and only if both $\frac{P_{i+1}P_{i+2}}{P_iP_{i+1}}$ and $\frac{P_iP_{i+1}}{P_{i+1}P_{i+2}}$ are bounded. When the first ratio is unbounded the distortion of this distance by f is unbounded near a; while if the second ratio is unbounded the distortion of distance by its inverse f^{-1} is unbounded near b.

However f is bounded and Lipschitz with respect to the additive distance defined by any $d : (a,b) \to \mathbb{R}$ which increases linearly over each P_iP_{i+1} such that $\frac{d(P_{i+2})-d(P_{i+1})}{d(P_{i+1})-d(P_i)}$ and its reciprocal are both bounded. So *any such d is* not *an alif or $\aleph$-distance on the interval.* For instance let d increase by the same amount say 1 on each segment P_iP_{i+1}, then the image of d is all of $\mathbb{R}$; or take $d(P_0) = 0$ and d increasing linearly by $1/2^{|i|}$ on each segment P_iP_{i+1}, then the image of d is only $(-1,2)$; etc. This last enables similar examples focused on *any* nonempty subinterval (a_k, b_k) of (a,b).

We note $\frac{d(P_{i+2})-d(P_{i+1})}{d(P_{i+1})-d(P_i)}$ bounded and $\frac{P_iP_{i+1}}{P_{i+1}P_{i+2}}$ unbounded implies that their product $\frac{d(P_{i+2})-d(P_{i+1})}{d(P_{i+1})-d(P_i)} \times \frac{P_iP_{i+1}}{P_{i+1}P_{i+2}}$ is unbounded. Therefore, *for any such d either the ratio of successive slopes $\frac{d(P_{i+2})-d(P_{i+1})}{P_{i+1}P_{i+2}} \div \frac{d(P_{i+1})-d(P_i)}{P_iP_{i+1}}$ or its reciprocal*

$\frac{d(P_{i+1})-d(P_i)}{P_iP_{i+1}} \div \frac{d(P_{i+2})-d(P_{i+1})}{P_{i+1}P_{i+2}}$ *is unbounded.* □

This suggests that maybe, in *any* dimension $n \geq 1$, for d to be an alif distance the local condition *aira* below is necessary and sufficient.

(ऐ) *The ratio* $\frac{d(P,Q)}{PQ} \div \frac{d(Q,R)}{QR}$ *of conversion ratios is bounded for all nearby adjacent pairs of elements* $\angle PQR := (PQ, QR)$; where 'nearby' means not only at all Q in a small neighbourhood of any $O \in U$, but also at all Q near any point O of ∂U. Thus this condition like aliph is also not symmetric in the two metrics: the distance d is defined only on U and may blow up near ∂U.

Before confirmng our hunch, we return to the case $n = 1$ – when *bends* $\angle PQR$ are necessarily one small segment PQ followed by another small segment QR of the same straight line – and look at some more

Examples (contd). The infinite series $\Sigma 1/2^n$ and $\Sigma 1/n^2$ are both convergent, the second converging much more slowly (it fails the ratio test). So we can choose that doubly infinite increasing sequence $P_i, i \in \mathbb{Z}$ of points of (a, b) to be such that near a the segments P_iP_{i+1} have length alternately 2^i and $1/i^2$. So half the time $\frac{P_{i+1}P_{i+2}}{P_iP_{i+1}}$ is becoming arbitrarily big and the other half arbitrarily small as $i \to -\infty$. This because $2^n/n^2 \to \infty$ as $n \to \infty$. Hence the distortion of Euclidean distance by the homeomorphism f of (a, b) which maps each P_iP_{i+1} linearly onto $P_{i+1}P_{i+2}$ is unbounded near a. On the other hand the distortion of the additive distance d which increases linearly over every P_iP_{i+1} by $1/i^2$ is bounded by 2. So d is not an alif distance:

Note *the graph of this non-aliph additive distance is zigzag,* with zags all of slope 1 near a, while the slopes of the zigs decreases to zero. So its forward blue cone at a consists of all rays having slopes in $[m, M]$ where $m = 0, M = 1$. Also, one can check that no ray with positive slope cuts it again sufficiently close to a, so d is right differentiable at a with right derivative zero.

Likewise the distortion of a d increasing linearly over every P_iP_{i+1} by 2^i is bounded by 2, so it too is not alif, but its forward blue cone at a consists of all rays having slopes $m = 1$ through $M = \infty$, i.e., the y-axis.

The condition (ऐ) can also fail thus: at all points of U the forward and backward cones are bounded away from the axes, but as we approach a boundary point, the ratio $M/m \to \infty$. Such a distance d is also not alif. □

Theorem. *A continuous distance d on U is alif iff it is aira.*

Proof. We are given a homeomorphism f of $\overline{U}$ identity on ∂U such that $\frac{d(f(P),f(Q)))}{d(P,Q)}$ and its reciprocal are uniformly bounded by a finite constant for all distinct pairs $\{P, Q\}$ of points of U. Indeed we'll need this condition only for pairs that are sufficiently close to each other in the Euclidean distance of the convex bounded open subset U of $\mathbb{R}^n$. (More generally, if $\overline{U} \subset \mathbb{R}^n$ is a compact manifold-with-boundary the same result holds if we use minimum arc length in U, on small segments of U this coincides with Euclidean distance.)

Since $\overline{U}$ is compact our job reduces to showing under condition (ऐ) on d that f is Lipschitz with respect to the Euclidean distance near any $O \in \overline{U}$. If $O \in U$ is not a fixed point of f we can compose with a quasi d-isometry of

U, identity on ∂U, which brings $f(O)$ back to O. Say, a diffeomorphism of U onto itself, identity outside a neighbourhood of $Of(O)$, which translates a neighbourhood of $f(O)$ back along this segment to one of O. (More generally, take any arc in U from $f(O)$ to O tangent to an apt smooth vector field zero outside a neighbourhood of this arc, to obtain a diffeomorphism of U onto itself, identity outside this neighbourhood, which translates a neighbourhood of $f(O)$ along this arc to one of O.) Its first derivatives being bounded in distance d as well, this gives us the required quasi-isometry. So we can assume without loss of generality that $O \in U$ is a fixed point of f. In case O is in the interior of the closed subset $F \supseteq \partial U$ of all fixed points of f there is nothing to do. Otherwise, if f were not Lipschitz in a neighbourhood of O, it would follow, using the fact that f is Lipschitz here with respect to d, that near O there are points $P, Q = f(P), R = f(Q)$ for which $\frac{d(P,Q)}{PQ} \div \frac{d(Q,R)}{QR}$ or its reciprocal is arbitrarily big, contradicting condition (ਅ).

The other direction follows because otherwise, as in above examples, we can find a small arc $\widetilde{OA}$ emanating strictly away from any $O \in \overline{U}$ at which d is not aira, together with a non-Lipschitz homeomorphism f of $\widetilde{OA}$ keeping O and A fixed, but which is Lipschitz with respect to the restriction of d. Further we can, e.g., by rotating the arc if $O \in U$, extend f to a homeomorphism of $\overline{U}$ keeping ∂U fixed. (We note that this arc $\widetilde{OA}$ can be very wavy: for example, if d is smooth, and doesn't grow too steeply if we approach the boundary normally, then aira can fail only for an $O \in \partial U$, and that only if—so this can happen only *for* $n \geq 2$—we approach O along a non-normal arc.) □

The first letter of all Semitic scripts, e.g., Hebrew and Arabic, as well as of Shahmukhi, the other main script in which Punjabi is written, is called *alif.* It is equivalent to the second letter *aira* ਅ of the Gurmukhi alphabet. After HGH, the case, hyperbolic distance on an open ball, was discussed in appendix alif (ℵ) of Siebenmann and Sullivan's *On Complexes that are Lipschitz Manifolds* (CLM), from the same 1977 conference. Note 22 cooked a whole lot, for example the distances of Notes 12 and 14 are not alif. The thunch that property alif is equivalent to property aira belongs to the on-going salvage operation underway since then … However in hindsight (ਅ) was all but there in Note 13 where it is observed that the ratio of slopes $\frac{s_{i+1}}{s_i}$ should be bounded. Notes 12 and 14 failed to take into account that, *for* $n \geq 2$, when slopes are in transverse directions, this ratio is not bounded for those distances either.

K. S. Sarkaria

P.S.– More generally an ਅ-distance on M^n satisfies $\frac{d(P,Q)}{PQ} \div \frac{d(Q,R)}{QR}$ bounded in suitable local coordinates which we can call an ℵ-structure. A Riemannian metric implies a smooth structure, and Steenrod showed any smooth manifold admits one. Likewise, any ℵ-manifold admits an ਅ-distance, which might help in understanding the obstruction to this structure in dimension $n = 4$. Also, very pertinent here is his remark that a mathematician "must live with disorganized facts imbedded in a sea of mud" till light hopefully dawns …