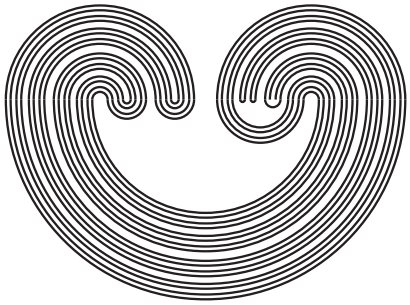
Topology Proceedings



Web: http://topology.auburn.edu/tp/
Mail: Topology Proceedings
Department of Mathematics & Statistics
Auburn University, Alabama 36849, USA
E-mail: topolog@auburn.edu
ISSN: 0146-4124

COPYRIGHT © by Topology Proceedings. All rights reserved.

TOPOLOGY PROCEEDINGS
Volume 30, No. 1, 2006
Pages 129-152

METRICS ON HYPERSPACES: A PRACTICAL APPROACH

WŁODZIMIERZ J. CHARATONIK AND MATT INSALL

Abstract. We present several new distance functions on hyperspaces and investigate their properties, relative to some natural application considerations.

1. Motivation

Consider a robot designed to sculpt. At first, it has a piece of wood (in mathematical terms a solid region in $\mathbb{R}^3$), and its goal is to produce a certain shape (another solid in $\mathbb{R}^3$), imbedded in the original one, by removal of the excess (see Figure 1).

We wish to measure the robot's progress, using a distance function on the hyperspace of closed subsets of $\mathbb{R}^3$. If the first cut is as depicted in Figure 1, then the Hausdorff distance between the unfinished sculpture and the goal remains unchanged, in spite of the fact that progress has definitely been made. This suggests to us that the Hausdorff metric is not the right distance function to use for this purpose. Some of the distance functions we will define here will be much better suited for such purposes.

2000 *Mathematics Subject Classification.* 54B20, 54A20, 54E99, 28A25, 54E35.

Key words and phrases. accurate distance function, discrepancy function, distance function, hyperspace, metric, pseudo-metric, quasimetric, symmetric, Whitney map.

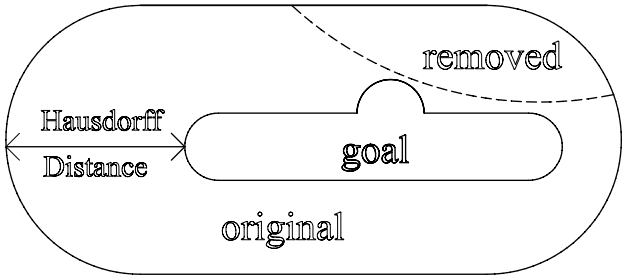


FIGURE 1

2. Examples of distances defined by discrepancy functions

In [2], we introduced the notion of a discrepancy function, and showed how this new notion can be used to define new distance functions on hyperspaces. In this section of the present work, we discuss several examples that are defined in this manner. Later, we will consider another new way to define distances between sets.

2.1 Preliminary terminology

For our purposes, a *distance function* on a set X is a function $d : X \times X \to [0, \infty]$. That is, d is a nonnegative extended real-valued bivariate function on X. We will focus on distance functions that satisfy the following properties:

$$d(x, x) = 0,$$

$$d(x, y) = d(y, x),$$

and

$$d(x, y) = 0 = d(y, x) \text{ implies } x = y.$$

Such a distance function is called a *symmetric*. (See [3].) Of course, a *metric* is a distance function that satisfies the above property, as well as the *triangle inequality*

$$d(x, z) \leq d(x, y) + d(y, z).$$

Some of the distance functions we shall consider are not metrics, but we shall compare them to each other and to various metrics on hyperspaces.

In working with symmetrics, we will find it convenient to consider intermediate distance functions that are not symmetrics. In particular, if a distance function is only known to satisfy $d(x,x) = 0$, we emphasize the fact that it may not satisfy $d(x,y) = 0 = d(y,x)$ implies $x = y$ by referring to d as a *pseudo-symmetric*. But if d is a pseudo-symmetric that also satisfies the triangle inequality, then we say that d is a *pseudo-metric*.

Another useful type of distance function we will discuss is a *quasi-symmetric*. This is a distance function that definitely satisfies $d(x,y) = 0 = d(y,x)$ implies $x = y$, but perhaps no other properties of a metric. A *quasi-metric* is a quasi-symmetric that satisfies the triangle inequality.

2.2 The distance functions

We will work in $\mathbb{R}^2$, although much of our work here is easily generalized to a larger class of metric spaces (see, for example, [2]). Let X be a compactum in $\mathbb{R}^2$, and let 2^X denote the collection of nonempty, closed (hence compact) subsets of X. For $A \subset X$ and any function $r : A \to [0,\infty]$, set

$$\mathcal{N}(A,r) = \bigcup_{x\in A} \mathcal{B}(x, r(x)),$$

where for $x \in X$ and $\varepsilon \in [0,\infty]$,

$$\mathcal{B}(x,\varepsilon) = \{y \in X | \|y - x\| < \varepsilon\}.$$

For $A, B \in 2^X$, we set

$$s_H(A,B) = \inf\{\varepsilon \in [0,\infty] | B \subset \mathcal{N}(A,\varepsilon)\},$$

and

$$s_{HL}(A,B) =$$

$$\left(\frac{1}{\pi}\left[\inf\{\lambda(\mathcal{N}(A,r)) | r : A \to (0,\infty], \text{ and } B \subset \mathcal{N}(A,r)\} - \lambda(A)\right]\right)^{\frac{1}{2}}$$

where λ denotes the Lebesgue measure. The division by π and then the application of the square root function, in the definition

of s_{HL}, make the resulting distance function agree with the usual distance function on the singletons, a very natural and convenient convention. Moreover, we know that without the square root, the result would definitely not be a quasi-metric. To see this, consider the points $x_0 = (0,0)$, $x_1 = (1,0)$, and $x_2 = (2,0)$. Then the s^2_{HL} distance between $\{x_0\}$ and $\{x_1\}$ and the s^2_{HL} distance between $\{x_1\}$ and $\{x_2\}$ are equal to one, but the s^2_{HL} distance between $\{x_0\}$ and $\{x_2\}$ is 4. (Also, see [2], and see our conjecture below.) This convention generalizes neatly to $\mathbb{R}^n$, because in that case, we would divide by the Lebesgue measure of the unit ball in $\mathbb{R}^n$, and then apply the n^{th} root to the result. Next, let μ be the *traditional* Whitney map (see [4, Exercise 13.5, p. 108], i.e., we mean to use the Whitney map defined using distances between points, which is invariant under isometries), and for $A, B \in 2^X$, set

$$s_W(A,B) = \mu(A \cup B) - \mu(A).$$

The mappings s_H and s_W are quasi-metrics and can be used to define metrics on 2^X (see [2, examples 2.14 and 2.13] and [2, Theorem 2.1]). The mapping s_{HL} is a quasi-symmetric, but we conjecture that it is a quasi-metric. In fact, here we consider six such distance functions, each of which is defined from its corresponding quasi-(sym)metric as follows: Given a quasi-(sym)metric s, define D by

$$D(A,B) = \max\{s(A,B), s(B,A)\},$$

or by

$$D(A,B) = s(A,B) + s(B,A).$$

We denote the resulting six distance functions as follows:

(1) $H_{\max}$ is the metric defined as above using max and the quasi-metric s_H:

$$H_{\max}(A,B) = \max\{s_H(A,B), s_H(B,A)\}.$$

(Note: This is the Hausdorff metric, as in [4, Definition 2.1, p. 11].)

(2) H_+ is the metric defined as above using addition and the quasi-metric s_H:

$$H_+(A,B) = s_H(A,B) + s_H(B,A).$$

(3) $HL_{\max}$ is the distance function defined as above using max and the quasi-symmetric s_{HL}:

$$HL_{\max}(A,B) = \max\{s_{HL}(A,B), s_{HL}(B,A)\}.$$

(Note: We call this the Hausdorff-Lebesgue symmetric.)

(4) HL_+ is the symmetric defined as above using addition and the quasi-symmetric s_{HL}:

$$HL_+(A,B) = s_{HL}(A,B) + s_{HL}(B,A).$$

(5) $W_{\max}$ is the metric defined as above using max and the quasi-metric s_W:

$$W_{\max}(A,B) = \max\{s_W(A,B), s_W(B,A)\}.$$

(6) W_+ is the metric defined as above using addition and the quasi-metric s_W:

$$W_+(A,B) = s_W(A,B) + s_W(B,A).$$

We will show, using some simple examples, that of these 6 distance functions, W_+ is arguably the best for the type of applications we described previously, but we will also show that all of these are computationally expensive.

In the sequel, when we use the symbol $D_{\max}$ for a metric, we mean $D_{\max} \in \{H_{\max}, HL_{\max}, W_{\max}\}$, and when we use the symbol D_+, we mean $D_+ \in \{H_+, HL_+, W_+\}$. The symbol D can stand for any of the following symbols: H, HL, W, $H_{\max}$, H_+, $HL_{\max}$, HL_+, $W_{\max}, W_+$. To deal with computational complexity, we will introduce later a variation on the HL symmetrics which leads to polynomial time approximations.

2.3 Using discrepancy functions to define the distance functions

For each $A \in 2^X$, let

$$\mathcal{F}_H(A) = \{\mathcal{N}(A,\varepsilon) | \varepsilon \in [0,\infty]\}.$$

Also, let

$$\mathcal{D}(\mathcal{F}_H) = \{(A,U) | U \in \mathcal{F}_H(A)\},$$

and define the following extended real-valued function on $\mathcal{D}(\mathcal{F}_H)$,

$$\omega_H(A,U) = \sup\{d(x,A) | x \in U\},$$

where d is the usual Euclidean distance function in the plane. Next, let

$$\mathcal{F}_{HL}(A) = \{\mathcal{N}(A,r) | r : A \to [0,\infty]\},$$

and set

$$\mathcal{D}(\mathcal{F}_{HL}) = \{(A,U) | U \in \mathcal{F}_{HL}(A)\}.$$

Define, for $(A,U) \in \mathcal{D}(\mathcal{F}_{HL})$,

$$\omega_{HL}(A,U) = \left[\frac{1}{\pi}\left(\lambda(U) - \lambda(A)\right)\right]^{\frac{1}{2}},$$

where λ is the Lebesgue measure. Finally, set

$$\mathcal{F}_W(A) = \{U \subset X | A \subset U = \bar{U}\},$$

$$\mathcal{D}(\mathcal{F}_W) = \{(A,U) | U \in \mathcal{F}_W(A)\},$$

and, for $(A,U) \in \mathcal{D}(\mathcal{F}_W)$,

$$\omega_W(A,U) = \mu(U) - \mu(A),$$

where μ is the *traditional* Whitney map. Each $\omega \in \{\omega_H, \omega_W\}$ is an example of a *proper t-discrepancy function* (see [2]). This means that each satisfies the following, for $A, B, C \in 2^X$ and for $D \in \{H, W\}$:

(df1) $\inf\{\omega(A,U) | U \in \mathcal{F}_D(A)\} = 0$.

(df2) If $A \subset B \subset C$, $U \in \mathcal{F}_D(A)$, $V \in \mathcal{F}_D(B)$, $B \subset U$, and $C \subset V$, then there is $W \in \mathcal{F}_D(A)$ such that $C \subset W$ and $\omega(A,U) + \omega(B,V) \geq \omega(A,W)$.

(df3) If $A \subsetneq B$, then $\inf\{\omega(A,U) | B \subset U\} > 0$.

Moreover, for each such discrepancy function, we may define a quasimetric s on 2^X (see [2]) via the formula

$$s(A,B) = \inf\{\omega(A,U) | B \subset U\}.$$

In fact, this provides a way to define the quasi-symmetrics we considered for the six symmetrics we studied previously:

(1) $s_H(A,B) = \inf\{\omega_H(A,U) | B \subset U \in \mathcal{F}_H(A)\}$,
(2) $s_{HL}(A,B) = \inf\{\omega_{HL}(A,U) | B \subset U \in \mathcal{F}_{HL}(A)\}$,
(3) $s_W(A,B) = \inf\{\omega_W(A,U) | B \subset U \in \mathcal{F}_W(A)\}$,

Our conjecture that s_{HL} is a quasi-metric (see [2, Theorem 2.1]) amounts to the claim that ω_{HL} is a proper t-discrepancy function.

2.4 Comparisons of the six chosen symmetrics

In figures 2, 3, and 4, we present a set with 2 approximations. The approximated set S is represented by a bent band and may represent a cross-section of a terrain with one hill. The approximation A to S is represented by the solid line and may represent an approximate terrain with one error. The other approximation, B, is represented by the dashed line and has another error, in addition to the error of A. Thus, intuitively, A is always a better approximation to S than is B. Note that although the set S is represented in our figures by a band, we intend S to be only a one-dimensional subset of $\mathbb{R}^2$. In fact, we may take S to be the polygonal arc with vertices $(0,0)$, $(2,0)$, $(3,1)$, $(4,0)$, and $(6,0)$.

In Figure 2, a terrain S with one hill is approximated first by a terrain A with one hill and one flagpole and then by a terrain B with one hill and two flagpoles. The flagpoles represent errors in the approximations. The distances

$$H_{\max}(A,S), H_{\max}(B,S), H_+(A,S), \text{ and } H_+(B,S)$$

satisfy

$$H_{\max}(A,S) = H_{\max}(B,S), \text{ and } H_+(A,S) = H_+(B,S).$$

Thus, both types of "Hausdorff" distances fail to recognize the difference in the level of error of A versus B (as approximations to S). However, we have

$$HL_{\max}(A,S) < HL_{\max}(B,S), HL_+(A,S) < HL_+(B,S),$$

$$W_{\max}(A,S) < W_{\max}(B,S), \text{ and } W_+(A,S) < W_+(B,S),$$

so all our other new symmetrics serve better to distinguish A from B as a better approximation to S.

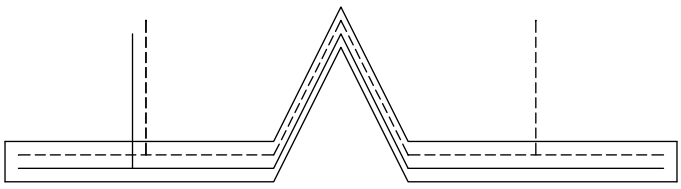


Figure 2

Figure 3 shows the same terrain S with two approximants, A, which is again as in Figure 2, and B, which now has a smaller second error very close to the first error. In this case, we demonstrate that all of the symmetrics $H_{\max}$, H_+, $HL_{\max}$, and HL_+ can fail simultaneously to distinguish the better approximation A from the worse approximation B. On the other hand, $W_{\max}$ and W_+ do distinguish these approximants.

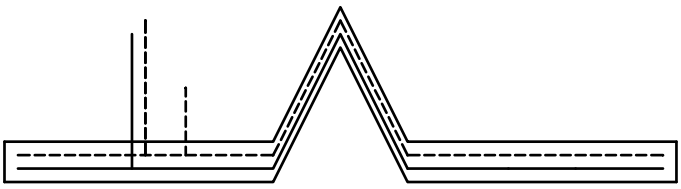


FIGURE 3

In Figure 4, the approximant A is the same as in Figure 3 (S is a proper subset of A), while B is a proper subset of A as well. But B does not compare to S, and the length of the interval $S \setminus B$ is sufficiently small. This example shows that all the D_+ symmetrics can distinguish A from B (as approximations to S), while, for this example, none of the $D_{\max}$ symmetrics agrees with intuition. To see this, observe that in each case, if ω is the corresponding discrepancy function, then

$$D_{\max}(S, A) = s_\omega(S, A),$$
$$D_{\max}(S, B) = \max\{s_\omega(S, A), s_\omega(A, B)\} = s_\omega(S, A),$$
$$D_+(S, A) = s_\omega(S, A),$$

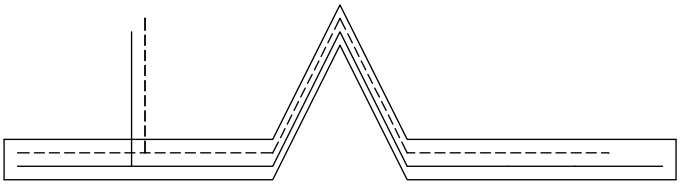


FIGURE 4

and

$$D_+(S,B) = s_\omega(S,A) + s_\omega(B,A) > s_\omega(S,A) = D_+(S,A).$$

In Table 1, we list the figure numbers and summarize the preceding discussions. For a symmetric D and a given figure, the corresponding entry in the table is a "+" if that example demonstrates

$$D(A,S) < D(B,S),$$

while a "–" in the table indicates

$$D(A,S) = D(B,S).$$

Figure	$H_{\max}$	H_+	$HL_{\max}$	HL_+	$W_{\max}$	W_+
2	–	–	+	+	+	+
3	–	–	–	–	+	+
4	–	+	–	+	–	+

Table 1

2.5 Some implications

The following proposition is easily deduced and applies to figures 2 and 3.

Proposition 2.1. *If $S \subset A \subset B$, then*

(1) *if $D \in \{H, HL, W\}$, then $D_{\max}(A,S) = D_+(A,S)$ and $D_{\max}(B,S) = D_+(B,S)$;*
(2) *$H_{\max}(A,S) < H_{\max}(B,S)$ implies that $HL_{\max}(A,S) < HL_{\max}(B,S)$ and $HL_{\max}(A,S) < HL_{\max}(B,S)$ implies $W_{\max}(A,S) < W_{\max}(B,S)$;*
(3) *if $A \neq B$, then $W_{\max}(A,S) < W_{\max}(B,S)$, even if $H_{\max}(A,S) = H_{\max}(B,S)$ or $HL_{\max}(A,S) = HL_{\max}(B,S)$.*
(4) *There are some cases in which $A \neq B$ but $H_{\max}(A,S) = H_{\max}(B,S)$ and $HL_{\max}(A,S) = HL_{\max}(B,S)$ (see Figure 3).*

The implications described in the following result are applicable to the situation depicted in Figure 4.

Theorem 2.1. *If $A = S \cup B$, then for $D \in \{H_{\max}, H_+, HL_{\max}, HL_+, W_{\max}, W_+\}$, we have*

$$D(A,S) \leq D(B,S),$$

and if moreover $A \neq B$, then for $D \in \{H, HL, W\}$, we have

$$D_+(A,S) < D_+(B,S),$$

but it is possible that, under these circumstances,

$$D_{\max}(A,S) = D_{\max}(B,S).$$

The proof of this is left as an exercise.

2.6 Absorbing distance functions

Let $\mathcal{H}$ be a hyperspace (i.e., a subset of 2^X), and let D be a distance function defined on $\mathcal{H}$. Then we say that D is *absorbing* provided that it satisfies

$$D(A,B) \geq D(A \cup C, B \cup C),$$

whenever $A, B, C, A \cup C, B \cup C \in \mathcal{H}$. A motivation for this notion follows. Let $\{C_\alpha | \alpha \in [0,1]\}$ be an order arc in $\mathcal{H}$ (i.e., $C_\alpha \subseteq C_\beta$, for $\alpha \leq \beta$) with $C_0 \subset A \cap B$, for some $A, B \in \mathcal{H}$, and $C_1 = X$. Assume that for each $\alpha \in [0,1]$, the sets $C_\alpha \cup A$ and $C_\alpha \cup B$ are in $\mathcal{H}$ and consider the function $\varphi : [0,1] \to [0,1]$ that is defined by

$$\varphi(\alpha) = D(C_\alpha \cup A, C_\alpha \cup B).$$

We have $\varphi(0) = D(A,B)$ and $\varphi(1) = 0$. It is natural to expect that φ is monotone nonincreasing (i.e. for $\alpha \leq \beta$, $\varphi(\alpha) \geq \varphi(\beta)$). But this can only be the case if the symmetric D is an absorbing symmetric. Thus, requiring a symmetric D on $\mathcal{H}$ to be absorbing is natural. However, examples of non-absorbing metrics are easy to come by, as we show.

Example 2.1. Let X be a three-point set, say $X = \{p, q, r\}$. Label some points of the Euclidean plane with the subsets of X as follows:

$$(-1,0) \mapsto \{p\}, (1,0) \mapsto \{q\}, (2,0) \mapsto \{r\}, (-2,1) \mapsto \{p,q\},$$
$$(1,1) \mapsto \{q,r\}, (2,1) \mapsto \{p,r\}, (0,2) \mapsto \{p,q,r\}$$

For subsets A, B of X, define $D(A,B) =\| x - y \|$ if and only if $x \mapsto A$ and $y \mapsto B$. For $a, b \in X$, set $d(a,b) = D(\{a\}, \{b\})$. The metric space (X,d) then has as its hyperspace of closed sets the

power set of X, and on this hyperspace, D is a metric that is not absorbing, because $D(\{p,q\},\{p,r\}) > D(\{q\},\{r\})$.

Theorem 2.2. *Let* $D \in \{H_{\max}, H_+, HL_{\max}, HL_+, W_{\max}, W_+\}$. *Then* D *is absorbing.*

Proof: For $H_{\max}$, H_+, $HL_{\max}$, and HL_+, this is an immediate consequence of the definitions of the various distances. For $W_{\max}$ and W_+, the desired result follows from the property

$$A \subset B \text{ implies } \mu(B \cup C) - \mu(A \cup C) \leq \mu(B) - \mu(A),$$

which holds for the traditional Whitney map that we are using here (but not for all Whitney maps; see [1, Proposition 1]). □

2.7 Accurate distance functions

A distance function D on a hyperspace $\mathcal{H}$ is *accurate* if it satisfies

$$A \subset B \subsetneq C \text{ implies } D(A,B) < D(A,C).$$

Although we showed in the preceding section that all six of the symmetrics we consider here are absorbing, we now will show that they are not all accurate.

Theorem 2.3. *Let* $D \in \{H_{\max}, H_+, HL_{\max}, HL_+, W_{\max}, W_+\}$. *Then* D *is accurate if and only if* $D \in \{W_{\max}, W_+\}$.

Proof: Figure 3 shows that if $D \in \{H_{\max}, H_+, HL_{\max}, HL_+\}$, then D is not accurate. The desired result for $W_{\max}$ and W_+ follows from the properties of the Whitney map μ, because if $A \subset B \subsetneq C$, then

$$\begin{aligned} &W_{\max}(A,B) = W_+(A,B) = \mu(B) - \mu(A) \\ &< \mu(C) - \mu(A) = W_{\max}(A,C) = W_+(A,C). \end{aligned}$$ □

2.8 Computational considerations

Here, we explain how to compute approximate distances between sets using a computer, and we discuss some related computational considerations.

First, we establish some notation for our purposes. Specifically, when we write $f \in \mathcal{O}(g)$, we mean that f and g are functions

of one or two natural number variable(s) m (and n), such that there exist $c_0, M > 0$ for which if $m > M$ (and $n > M$), then $f(m,n) < c_0 g(m,n)$. We will assume that the sets A and B are finite sets in a finite subspace X of $\mathbb{R}^2$, say $|A| = n$ and $|B| = m$. Thus, let

$$A = \{a_1, ..., a_n\} \text{ and } B = \{b_1, ..., b_m\}.$$

For our metrics, we first need to compute quasi-symmetrics. (See [2, Theorem 2.1] for more detailed discussion of the relationship between our quasi-symmetrics and our metrics.) This is quite simple for $H_{\max}$ and H_+. In fact, for these two, the quasi-symmetrics are the same, and we denote them both by s_H:

$$s_H(A,B) = \max_{1\le j\le m} \min_{1\le i\le n} d(a_i, b_j),$$

and, of course,

$$s_H(B,A) = \max_{1\le i\le n} \min_{1\le j\le m} d(a_i, b_j).$$

Then for our two H metrics, we have

$$H_{\max}(A,B) = \max\{s_H(A,B), s_H(B,A)\},$$

and

$$H_+(A,B) = s_H(A,B) + s_H(B,A).$$

It follows that computation of these two distances is of quadratic complexity, because mn comparisons are to be made in computing $s_H(A,B)$.

To compute the W metrics, we need to calculate, for instance, $\mu(A)$ where μ is the Whitney map we described previously. This is done via the formula

$$\mu(A) = \sum_{i=2}^{\infty} 2^{-i} \mu_i(A),$$

where for each i,

$$\mu_i(A) = \max_{\substack{F \subset A \\ |F| = i}} \{\min\{d(a,b) | a, b \in F, a \neq b\}\}.$$

Moreover, we use the convention that $\max \emptyset = 0$, so that if $i > n = |A|$, then $\mu_i(A) = 0$. This allows us to truncate the series in the

computation of $\mu(A)$:

$$\mu(A) \approx \sum_{i=2}^{n} 2^{-i}\mu_i(A).$$

It follows that the complexity of computing $\mu(A)$ is $\mathcal{O}(2^n)$, since

$$\sum_{i=0}^{n} \binom{n}{i} = 2^n.$$

Now,

$$s_W(A,B) = \mu(A \cup B) - \mu(A),$$

and similarly,

$$s_W(B,A) = \mu(A \cup B) - \mu(B),$$

and then,

$$W_{\max}(A,B) = \max\{s_W(A,B), s_W(B,A)\},$$

and

$$W_+(A,B) = s_W(A,B) + s_W(B,A);$$

it follows that the complexity of these two metrics is $\mathcal{O}(2^{m+n})$. However, if $k < m+n$ and we approximate by truncating at the term $2^{-k}\mu_k(\cdot)$, i.e.,

$$\mu(A) = \sum_{i=2}^{\infty} 2^{-i}\mu_i(A) \approx \sum_{i=2}^{k} 2^{-i}\mu_i(A),$$

then the complexity for the resulting approximations is $\mathcal{O}((m+n)^k)$.

To compute the complexity of $HL_{\max}$ and HL_+, observe that in the definition of s_{HL}, we consider *all* functions $r : A \to [0,\infty)$, and so direct computation leads to exponential complexity.

Thus, let us describe a modification, HLM (HL from Hausdorff-Lebesgue, and M for "modified"), of the HL symmetrics that leads to polynomial time approximation complexity and still enjoys the benefits we outlined for $HL_{\max}$ and HL_+. Specifically, the $HLM_{\max}$ distance function will still distinguish between the approximants A and B illustrated in Figure 2, while the HLM_+ distance function will still distinguish the cases illustrated in figures 2 and 4. Let $A, B \in 2^X$. We compute a quasi-symmetric $s_{HLM}(A,B)$ according to the following recipe: Let $\mathcal{B}_C(p,\delta)$ denote the closed ball centered at the point p, and of radius δ. Next, let F be a function that assigns to a point b of B the set of all

closest points of A to b. Next, let $\rho(b) = \bigcup_{a\in F(b)} \mathcal{B}_C(a, d(b,a))$, and then let $R(A,B) = A \cup \left(\bigcup_{b\in B} \rho(b)\right)$. Then $s_{HLM}(A,B)$ is $\left[\frac{1}{\pi}\left(\lambda(R(A,B)) - \lambda(A)\right)\right]^{\frac{1}{2}}$. Thus, $s_{HLM}(A,B)$ is well-defined, regardless of the finitude of either A or B. In case $A = \{a_1, ..., a_m\}$ and $B = \{b_1, ..., b_n\}$ are finite (or in case they are approximated by finite sets), finding the closest points of A to points of B requires mn comparisons, and the calculation of areas may be implemented by imagining the points of X as locations in a matrix, and we "turn on" the points of $\rho(b_j)$ by replacing zeros with ones in the matrix. Specifically, let $Y \subset \mathbb{R}^2$ be a rectangular region (which we shall approximate with a uniform grid); let k, ℓ be positive integers; and, for $i \le k$, $j \le \ell$, let $(p_i, q_i) \in Y$ (so that the set $\{(p_i, q_j) | i \le k, j \le \ell\}$ is a finite approximation to Y). For any subset $S \subset Y$, set

$$G_S = [g_{ij}], \text{ where } g_{ij} = \begin{cases} 1 & \text{if } (p_i, q_j) \in S \\ 0 & \text{otherwise} \end{cases}.$$

Using this general approach to model a finite set of points with ones and zeros in a matrix, we may proceed as follows.

When we have turned on the points of A and the points of $\rho(b_j)$ for every j (this requires no more than K operations, where K is the number of points in the space X), we will have ones in the matrix corresponding only to the points of $R(A,B)$. Finally, to estimate the measure (i.e., area) of the region $R(A,B)$, we merely count the ones in the matrix, and this again requires on the order of K computations. From this we subtract the number, m, of points of the set A. The value thus obtained, namely $\left[\frac{1}{\pi}\left(|R(A,B)| - n\right)\right]^{\frac{1}{2}}$, is proportional to $s_{HLM}(A,B)$. Overall, this requires on the order of K^2mn computations. A similar calculation yields $s_{HLM}(B,A)$. Then our HLM symmetrics are computed accordingly:

$$HLM_{\max}(A,B) = \max\{s_{HLM}(A,B), s_{HLM}(B,A)\}$$

and

$$HLM_{+}(A,B) = s_{HLM}(A,B) + s_{HLM}(B,A).$$

Thus, the complexity of calculation of the HLM symmetrics is still polynomial, being on the order of K^2mn.

As the following example shows, the HLM symmetrics are different from the HL symmetric, but one can argue either way which one is closer to intuition.

Example 2.2. Let $A = [0,2] \times \{1\}$, and $B = [0,2] \times \{0\}$. Then, $s_{HLM}(A,B) = \left(\frac{4+\pi}{\pi}\right)^{\frac{1}{2}}$, but $s_{HL}(A,B) = \left(\frac{2\pi}{\pi}\right)^{\frac{1}{2}} = \sqrt{2}$ (see Figure 5). Thus, $s_{HL}(A,B) < s_{HLM}(A,B)$. Let $C = \{(1,1)\}$. Then $s_{HL}(A,B) = s_{HLM}(C,B) = s_{HL}(C,B)$.

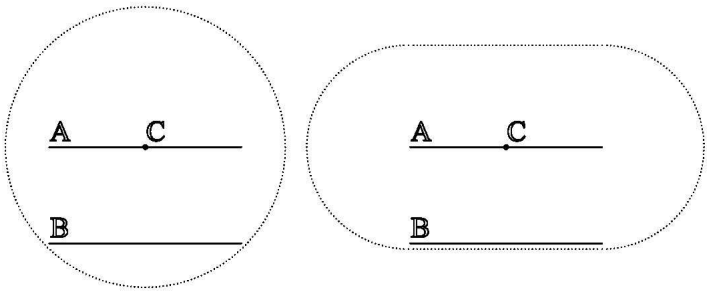


FIGURE 5

3. Integral metrics

The distance functions considered above were, in each case, either accurate (see subsection 3.2) or computationally feasible, but not both. Here, we present another construction of distance functions we call integral distance functions, because they are defined not by discrepancy functions but by certain integrals, and we show that in the cases in which they are metrics, these integral metrics are both accurate and computationally feasible (i.e., they can be reasonably approximated by algorithms that run in polynomial time).

Let (X,d) be a compact metric space on which there is given a finite measure, λ, for which all Borel sets are measurable. Given a closed set $A \subset X$, denote by $d_A : X \to [0,\infty)$ the function that assigns to a point $x \in X$ the distance from x to the set A:

$$d_A(x) = d(x,A) = \inf\{d(x,y) | y \in A\}.$$

Given a second closed set $B \subset X$, we set

$$D_p(A,B) = \left(\int_X |d_A(x) - d_B(x)|^p d\lambda(x) \right)^{\frac{1}{p}},$$

where $p \in (0, \infty)$. For $p = \infty$, we define $D_p = D_\infty$ using the essential supremum:

$$D_\infty(A, B) = \text{ess_sup}\{|d_A(x) - d_B(x)||x \in X\}.$$

3.1 Properties of D_p

Given the measure λ on the compact metric space (X, d) and given $p \in (0, \infty]$, we have the following result, whose proof is an observation that for closed sets A and B in X, the value of $D_p(A, B)$ is merely the $\mathcal{L}^p$ distance from (the measure-theoretic equivalence class of) d_A to d_B.

Theorem 3.1. *Given $p \in (0, \infty]$, the function D_p is a pseudo-symmetric. If $p \geq 1$, then D_p is a pseudo-metric.*

Thus, we call D_p an *integral pseudo-symmetric* on the hyperspace of X.

Now, recall that the measure λ is called a *strictly positive measure* provided that for each non-empty open set $U \subset X$, we have $\lambda(U) > 0$. This is a natural condition to impose on the measure λ, and it leads to the following result.

Theorem 3.2. *The integral pseudo-symmetrics D_p are symmetrics if and only if the given measure λ is strictly positive. Moreover, if λ is strictly positive, and if $p \geq 1$, then D_p is a metric.*

Observation 3.1. *If λ is strictly positive, then D_p is topologically equivalent to the Hausdorff metric H.*

Observation 3.2. *If λ is strictly positive, then*

$$D_\infty(A, B) = \max\{|d_A(x) - d_B(x)| \ |x \in X\}.$$

We will show in Theorem 3.5 that D_∞ is just the Hausdorff distance.

The metrics we are here calling integral metrics are accurate.

Theorem 3.3. *If the measure λ on X is strictly positive and if $1 \leq p < \infty$, then the integral metric D_p is an accurate metric.*

Proof: The proof of this is a bit like the proof of the second observation above. Let $A \subset B \subsetneq C$. We wish to show that $D_p(A, B) < D_p(A, C)$. Thus, let $c \in C \setminus B$, and denote by r the distance from c to the set B. Let U be the open ball centered

at c with radius $\frac{r}{2}$, so that $d_C(x) < \frac{r}{2} < d_B(x)$, for every $x \in U$. Since λ is strictly positive, we have $\lambda(U) > 0$. But also,

$$D_p(A, B)^p = \int_X (d_A(x) - d_B(x))^p d\lambda(x),$$

and

$$D_p(A, C)^p = \int_X (d_A(x) - d_C(x))^p d\lambda(x).$$

To see the inequality $D_p(A, B) < D_p(A, C)$, it is enough to observe that for all $x \in X$,

$$d_A(x) - d_B(x) \leq d_A(x) - d_C(x),$$

while strict inequality holds for $x \in U$. □

When are integral metrics absorbing? We answer this in the following result.

Theorem 3.4. *Let $p \in (0, \infty]$. Then D_p is an absorbing distance function.*

Proof: Let $A, B, C \subset X$. If $x \in X$, then suppose, without loss of generality, that $d_B(x) \geq d_A(x)$. If $d_C(x) \geq d_B(x)$, then

$$|d_{A\cup C}(x) - d_{B\cup C}(x)| = d_B(x) - d_A(x).$$

If $d_C(x) \leq d_A(x)$, then

$$|d_{A\cup C}(x) - d_{B\cup C}(x)| = d_C(x) - d_C(x) = 0.$$

If $d_A(x) \leq d_C(x) \leq d_B(x)$, then

$$\begin{aligned} |d_{A\cup C}(x) - d_{B\cup C}(x)| &= |d_A(x) - d_C(x)| \\ &= d_C(x) - d_A(x) \leq d_B(x) - d_A(x). \end{aligned}$$

Thus in all cases, we have

$$|d_{A\cup C}(x) - d_{B\cup C}(x)| \leq |d_A(x) - d_B(x)|,$$

and it follows that $D_p(A, B) \geq D_p(A \cup C, B \cup C)$, because

$$D_p(A, B) = \begin{cases} \left(\int_X |d_A(x) - d_B(x)|^p d\lambda(x)\right)^{\frac{1}{p}} & \text{if } p < \infty \\ \text{ess_sup}\{\, |d_A(x) - d_B(x)|\, |\, x \in X\} & \text{if } p = \infty \end{cases},$$

and

$$D_p(A \cup C, B \cup C)$$
$$= \begin{cases} \left(\int_X |d_{A\cup C}(x) - d_{B\cup C}(x)|^p d\lambda(x)\right)^{\frac{1}{p}} & \text{if } p < \infty \\ \text{ess_sup}\{\, |d_{A\cup C}(x) - d_{B\cup C}(x)| \,|\, x \in X\} & \text{if } p = \infty \end{cases}.$$

Thus, D_p is absorbing, as claimed. □

As a consequence of Theorem 3.3, when we consider the integral metrics applied to the cases exemplified by Figure 2 and Figure 3, we may add a column to Table 1 for integral metrics with plus signs in the first two rows. For the case of Figure 4, we calculate as follows, to show that a plus sign belongs in the third row also:

$$\begin{aligned} D_p(A,S)^p &= \int_X |d_A(x) - d_S(x)|^p d\lambda(x) \\ &= \int_L |d_A(x) - d_S(x)|^p d\lambda(x) + \int_R |d_A(x) - d_S(x)|^p d\lambda(x) \\ &= \int_L |d_A(x) - d_S(x)|^p d\lambda(x) \\ &= \int_L |d_B(x) - d_S(x)|^p d\lambda(x) \\ &< \int_L |d_B(x) - d_S(x)|^p d\lambda(x) + \int_R |d_B(x) - d_S(x)|^p d\lambda(x) \\ &= D_p(B,S)^p, \end{aligned}$$

where X is any compact rectangle, with sides parallel to the coordinate axes, containing the set $[0,6] \times [0,1]$, and L and R are closed rectangles, whose union is X, and whose intersection has empty interior, and which satisfy $B \subset L$ and $A \setminus B \subset R$.

We now wish to show that $D_\infty = H$ is the Hausdorff metric. First, we prove some related lemmas and a corollary.

Lemma 3.1. *Let $A = \{a\} \subset B \subset X$, where B is closed. Then*

$$D_\infty(A,B) \leq H(A,B).$$

Proof: Let $b \in B$ with $d(a,b) = H(A,B)$. If $x \in X$, let $x_B \in B$ with $d(x,x_B) = d_B(x)$. Then, by the triangle inequality,

$$|d_B(x) - d_A(x)| \leq d(x_B,a) \leq d(b,a) = H(A,B).$$

The desired result follows. □

Lemma 3.2. *Let $A \subset B \subset X$ with A, B closed. Then*

$$D_\infty(A, B) \leq H(A, B).$$

Proof: Let $Y = X/A$, and define $\rho : Y \times Y \to \mathbb{R}$ by the formula

$$\rho([x], [y]) = \min\{d(x, y), d_A(x) + d_A(y)\}.$$

We will show that ρ is a metric on Y. To see this, let $[x], [y] \in Y$. Then $\rho([x], [y]) = 0$ if and only if either $d(x, y) = 0$ or $d_A(x) + d_B(x) = 0$. In the first case, $x = y$, so that $[x] = [y]$. In the second case, $d_A(x) = 0$ and $d_A(y) = 0$, from which it follows that $[x], [y] \in A$ (since A is closed). Thus, $\rho([x], [y]) = 0$ if and only if $[x] = [y]$. Clearly, ρ is symmetric, so we turn to showing it satisfies the triangle inequality. Thus, let $[x], [y], [z] \in Y$ and suppose that

$$\rho([x], [y]) + \rho([y], [z]) \leq \rho([x], [z]).$$

We will consider three cases:

(1) $\rho([x], [y]) = d(x, y)$ and $\rho([y], [z]) = d(y, z)$;
(2) $\rho([x], [y]) = d_A(x) + d_A(y)$ and $\rho([y], [z]) = d(y, z)$;
(3) $\rho([x], [y]) = d_A(x) + d_A(y)$ and $\rho([y], [z]) = d_A(y) + d_A(z)$.

It is easy to see that from these three cases, the desired result follows in all eight of the possible cases.

Case 1. ($\rho([x], [y]) = d(x, y)$ and $\rho([y], [z]) = d(y, z)$.) In this case, we have

$$d(x, y) + d(y, z) \leq \rho([x], [z]) \leq d(x, z);$$

therefore, since the metric d satisfies the triangle inequality, we get

$$\rho([x], [y]) + \rho([y], [z]) = d(x, y) + d(y, z) = d(x, z) = \rho([x], [z]),$$

as desired.

Case 2. ($\rho([x], [y]) = d_A(x) + d_A(y)$ and $\rho([y], [z]) = d(y, z)$). Here, let $x_A, y_A \in A$ with $d_A(x) = d(x_A, x)$ and $d_A(y) = d(y_A, y)$. Then we have

$$d(x, x_A) + d(y_A, y) + d(y, z) \leq \rho([x], [z]) \leq d(x, z);$$

therefore, since $d(y_A, y) + d(y, z) \geq d(y_A, z) \geq d_A(z)$, it follows that

$$d(x, x_A) + d_A(z) \leq d(x, z),$$

i.e.,

$$d_A(x) + d_A(z) \leq d(x, z).$$

But then

$$\rho([x], [z]) = d_A(x) + d_A(z);$$

therefore, since also

$$d_A(x) + d_A(z) \leq d_A(x) + d_A(y) + d(y, z),$$

it follows that

$$\rho([x], [z]) \leq \rho([x], [y]) + \rho([y], [z]),$$

as desired.

Case 3. ($\rho([x], [y]) = d_A(x) + d_A(y)$ and $\rho([y], [z]) = d_A(y) + d_A(z)$.) Here, we have

$$d_A(x) + d_A(y) + d_A(y) + d_A(z) \leq \rho([x], [z]) \leq d_A(x) + d_A(z),$$

so that $d_A(y) = 0$, i.e., $y \in A$. Thus,

$$\begin{aligned} \rho([x], [y]) + \rho([y], [z]) &= d_A(x) + d_A(y) + d_A(y) + d_A(z) \\ &= d_A(x) + d_A(z) = \rho([x], [z]) \\ &\geq \rho([x], [y]) + \rho([y], [z]), \end{aligned}$$

so we are done showing that ρ is a metric.

Now we apply the previous lemma to the space Y. Fix $a \in A$ and $b \in B$ with $H(A, B) = d(a, b)$. Let $x \in X$, let $x_B \in B$ with $d(x, x_B) = d_B(x)$, and let $x_{BA} \in A$ with $\rho([x_B], A) = d(x_{BA}, x_B)$. Then

$$\begin{aligned} |d_B(x) - d_A(x)| &= |\rho_B(x) - \rho_{\{A\}}(x)| \\ &\leq \rho([x_B], A) = d(x_{BA}, x_B) \\ &\leq d(b, a) = H(A, B). \end{aligned}$$

The desired result follows. □

Corollary 3.1. *Let $p \in (0, \infty)$. Then for any closed $A, B \subset X$,*

$$D_p(A, B) \leq H(A, B)\lambda(X)^{\frac{1}{p}}.$$

Proof: The inequality shown in the last few lines of the proof of the previous lemma holds even when A is not contained in B and is useful here. For each $x \in X$, we have

$$|d_B(x) - d_A(x)| \leq H(A, B).$$

It follows that

$$\begin{aligned} D_p(A,B) &= \left(\int_X |d_B(x) - d_A(x)|^p d\lambda(x) \right)^{\frac{1}{p}} \\ &\leq \left(\int_X H(A,B)^p d\lambda(x) \right)^{\frac{1}{p}} \\ &= \left(H(A,B)^p \int_X d\lambda(x) \right)^{\frac{1}{p}} \\ &= H(A,B)\lambda(X)^{\frac{1}{p}}, \end{aligned}$$

so we are done. □

Now, our promised result relating D_∞ and H is immediate:

Theorem 3.5. *Let $A, B \subset X$ be closed. Then*

$$D_\infty(A,B) = H(A,B).$$

Proof: By the preceding lemmas, we have

$$D_\infty(A,B) \leq H(A,B).$$

Thus, we need demonstrate only the reverse inequality. Let $a \in A$ and $b \in B$ satisfy $H(A,B) = d(a,b)$. Then

$$H(A,B) = d(a,b) \leq \max\{d_A(b), d_B(a)\} \leq D_\infty(A,B),$$

so the desired result follows. □

3.2 Computational considerations for the integral metrics

In this subsection, we shall describe how to efficiently compute $D_p(A,B)$, given $p \in (0,\infty]$ and closed sets $A, B \subset X$. The algorithm we give runs in polynomial time. In fact, the following algorithm is simple: Approximate A, with a finite set, $\tilde{A}$, of cardinality m; approximate B with a finite set, $\tilde{B}$, of cardinality n; and approximate X with a finite set, $\tilde{X}$, of cardinality K. Then for a given point x of $\tilde{X}$, $d_A(x)$ is estimated by computing $d(x,\tilde{a})$, for each point $\tilde{a}$ of $\tilde{A}$, which requires m distance measurements. Similarly, estimation of $d_B(x)$ requires n distance measurements, so that estimation of $|d_A(x) - d_B(x)|$ is linear, being of the order

of $m+n$ calculations. Finally, since $\tilde{X}$ has K points, estimation of the integral

$$\int_X |d_A(x) - d_B(x)|^p d\lambda(x)$$

requires on the order of $K(m+n)$ calculations. Thus, approximation of $D_p(A,B)$ is of quadratic complexity.

3.3 Comparisons of D_p and D_q for $p, q \in (0, \infty]$

In this subsection, we let $p, q \in (0, \infty]$, and we consider two traditionally interesting cases: *(i)* $p < q$, and *(ii)* $\frac{1}{p} + \frac{1}{q} = 1$. First, recall from real analysis the inequality known as Hölder's inequality:

$$1 \leq p \leq q, \frac{1}{p} + \frac{1}{q} = 1, f \in \mathcal{L}^p, g \in \mathcal{L}^q \Longrightarrow \int_X |fg| d\lambda \leq \|f\|_p \|g\|_q.$$

Here, $\|f\|_p = \left(\int_X |f|^p d\lambda\right)^{\frac{1}{p}}$ and similarly for $\|g\|_q$. A simple consequence of Hölder's inequality is the following.

Lemma 3.3. *Let $1 \leq a \leq b$, $f \in \mathcal{L}^a$, $f \in \mathcal{L}^b$. Then*

$$\left(\frac{1}{\lambda(X)} \int_X |f|^a d\lambda\right)^{\frac{1}{a}} \leq \left(\frac{1}{\lambda(X)} \int_X |f|^b d\lambda\right)^{\frac{1}{b}}.$$

Proof: Let $g = 1$ and replace f by f^a. Let $p = \frac{b}{a}$, $q = (1 - \frac{a}{b})^{-1}$. Hölder's inequality implies

$$\int_X |f|^a d\lambda \leq \left(\int_X |f|^{ap} d\lambda\right)^{\frac{1}{p}} \lambda(X)^{1-\frac{a}{b}}.$$

Division by $\lambda(X)$, and then some rearrangement, yields

$$\frac{1}{\lambda(X)} \int_X |f|^a d\lambda \leq \frac{1}{\lambda(X)^{\frac{a}{b}}} \left(\int_X |f|^{ap} d\lambda\right)^{\frac{1}{p}} = \left(\frac{1}{\lambda(X)} \int_X |f|^b d\lambda\right)^{\frac{a}{b}},$$

from which the desired result follows. □

The authors are very grateful to Miron Bekker for pointing out this consequence of Hölder's inequality.

In our setting, the functions d_A, d_B are continuous, so the integrands are in both $\mathcal{L}^p$ and $\mathcal{L}^q$. Thus, we have the following result.

Theorem 3.6. *Let $A, B \subset X$ be closed, and $1 \leq p \leq q$. Then*

$$\begin{aligned}\frac{1}{\lambda(X)} D_1(A,B) &\leq \left(\frac{1}{\lambda(X)}\right)^{\frac{1}{p}} D_p(A,B) \\ &\leq \left(\frac{1}{\lambda(X)}\right)^{\frac{1}{q}} D_q(A,B) \\ &\leq H(A,B) = D_\infty(A,B).\end{aligned}$$

Proof: We prove that

$$\left(\frac{1}{\lambda(X)}\right)^{\frac{1}{p}} D_p(A,B) \leq \left(\frac{1}{\lambda(X)}\right)^{\frac{1}{q}} D_q(A,B).$$

The rest follows, of course. But, in fact, the desired result is an immediate consequence of the preceding lemma.

$$\left(\frac{1}{\lambda(X)}\right)^{\frac{1}{p}} D_p(A,B) = \left(\frac{1}{\lambda(X)} \int_X |d_B(x) - d_A(x)|^p d\lambda\right)^{\frac{1}{p}}$$

$$\leq \left(\frac{1}{\lambda(X)} \int_X |d_B(x) - d_A(x)|^q d\lambda\right)^{\frac{1}{q}} \leq \left(\frac{1}{\lambda(X)}\right)^{\frac{1}{q}} D_q(A,B). \quad \square$$

Now, we will give another consequence of Hölder's inequality, but first, let us point out that it is natural to consider what happens if we "normalize" our distance functions. Accordingly, we set

$$\begin{aligned}N_p(A,B) &= \left(\frac{1}{\lambda(X)} \int_X |d_A(x) - d_B(x)|^p d\lambda(x)\right)^{\frac{1}{p}} \\ &= \left(\frac{1}{\lambda(X)}\right)^{\frac{1}{P}} D_p(A,B),\end{aligned}$$

and, consequently, $N_\infty = D_\infty$. A reformulation of the preceding result in terms of N_p is, of course, the following.

Theorem 3.7. *Let $A, B \subset X$ be closed, and let $1 \leq p \leq q$. Then*

$$N_1(A,B) \leq N_p(A,B) \leq N_q(A,B) \leq N_\infty(A,B) = H(A,B).$$

But a direct application of Hölder's inequality gives us the following interesting result.

Theorem 3.8. *Let* $1 \leq p \leq q$ *with* $\dfrac{1}{p} + \dfrac{1}{q} = 1$*, and let* $A, B \subset X$ *be closed. Then*

$$(D_2(A,B))^2 \leq D_p(A,B) D_q(A,B),$$

and similarly,

$$(N_2(A,B))^2 \leq N_p(A,B) N_q(A,B).$$

Acknowledgment. The authors would like to express their gratitude to a referee for very careful reading of the preliminary manuscripts of this work and for very helpful comments.

References

[1] Włodzmierz J. Charatonik, *A metric on hyperspaces defined by Whitney maps*, Proc. Amer. Math. Soc. **94** (1985), no. 3, 535–538.

[2] W. J. Charatonik and M. Insall, *Metrics defined via discrepancy functions.* Preprint.

[3] Troy L. Hicks and B. E. Rhoades, *Fixed-points for pairs of mappings in d-complete topological spaces*, Internat. J. Math. Math. Sci., **16** (1993), no. 2, 259–266.

[4] Alejandro Illanes and Sam B. Nadler, Jr., *Hyperspaces: Fundamentals and Recent Advances.* Monographs and Textbooks in Pure and Applied Mathematics, 216. New York: Marcel Dekker, Inc., 1999.

(Charatonik) Department of Mathematics and Statistics; University of Missouri - Rolla; 1870 Miner Circle; Rolla, MO 65409-0020
E-mail address: wjcharat@umr.edu

(Insall) Department of Mathematics and Statistics; University of Missouri - Rolla; 1870 Miner Circle; Rolla, MO 65409-0020
E-mail address: insall@umr.edu