Dissimilarity Cumulation Theory and Subjective Metrics

Ehtibar N. Dzhafarov*   Hans Colonius

Purdue University   Oldenburg University

Abstract

We present a new mathematical notion, dissimilarity function, and based on it, a radical extension of Fechnerian Scaling, a theory dealing with the computation of subjective distances from pairwise discrimination probabilities. The new theory is applicable to all possible stimulus spaces subject to the following two assumptions: (A) that discrimination probabilities satisfy the Regular Minimality law; and (B) that the canonical psychometric increments of the first and second kind are dissimilarity functions. A dissimilarity function $D\mathbf{ab}$ for pairs of stimuli in a canonical representation is defined by the following properties: (1) $\mathbf{a} \neq \mathbf{b} \implies D\mathbf{ab} > 0$; (2) $D\mathbf{aa} = 0$; (3) If $D\mathbf{a}_n\mathbf{a}'_n \to 0$ and $D\mathbf{b}_n\mathbf{b}'_n \to 0$, then $D\mathbf{a}'_n\mathbf{b}'_n - D\mathbf{a}_n\mathbf{b}_n \to 0$; and (4) for any sequence $\{\mathbf{a}_n\mathbf{X}_n\mathbf{b}_n\}_{n\in\mathbb{N}}$, where $\mathbf{X}_n$ is a chain of stimuli, $D\mathbf{a}_n\mathbf{X}_n\mathbf{b}_n \to 0 \implies D\mathbf{a}_n\mathbf{b}_n \to 0$. The expression $D\mathbf{aXb}$ refers to the dissimilarity value cumulated along successive links of the chain $\mathbf{aXb}$. The subjective (Fechnerian) distance between $\mathbf{a}$ and $\mathbf{b}$ is defined as the infimum of $D\mathbf{aXb}+D\mathbf{bYa}$ across all possible chains $\mathbf{X}$ and $\mathbf{Y}$ inserted between $\mathbf{a}$ and $\mathbf{b}$.

KEYWORDS: deviation, dissimilarity, discrimination probability, Fechnerian Scaling, observation area, oriented distance, perceptual discrimination, Regular Minimality, same-different judgements, stimulus chains, stimulus space, subjective equality, symmetric distance, topology, uniformity

1.   Introduction

This paper introduces a new mathematical notion, *dissimilarity function*. This notion shares some properties with that of a *metric* (distance function), but it is considerably more general: in particular, dissimilarity generally satisfies neither the triangle inequality nor the symmetry constraint. Once defined on some space, however (e.g., space of stimuli), a dissimilarity function allows one to impose a metric on this space by means of the following *standard construction*: dissimilarity values are *cumulated* along all finite chains of points leading from point $\mathbf{a}$ to point $\mathbf{b}$, and the infimum of these cumulated values (which one can call the "lengths" of the chains leading from $\mathbf{a}$ to $\mathbf{b}$) is taken to be the *oriented* (asymmetric) distance from $\mathbf{a}$ to $\mathbf{b}$;[1] by adding

---

*Corresponding author. Department of Psychological Sciences, Purdue University, 703 Third Street, West Lafayette, IN 47907-2081, USA. E-mail address: ehtibar@purdue.edu (E.N. Dzhafarov).

[1] The oriented distance from $\mathbf{a}$ to $\mathbf{b}$ is a quantity $G\mathbf{ab} \geq 0$ which satisfies the triangle inequality (for any $\mathbf{x}$, $G\mathbf{ax}+G\mathbf{xb} \geq G\mathbf{ab}$) and the zero property ($G\mathbf{ab} = 0$ iff $\mathbf{a} = \mathbf{b}$), but which does not have to satisfy the symmetry constraint ($G\mathbf{ab} = G\mathbf{ba}$). If the latter is satisfied too, the distance is called conventional, or symmetric.

the oriented distance from **a** to **b** to that from **b** to **a** one gets a conventional (symmetric) distance between **a** and **b**. In view of this procedure we call the new mathematical theory to be presented the *Dissimilarity Cumulation* (DC) theory.

The motivation for this particular construction comes from its main application: computing "subjective" (*Fechnerian*) distances among stimuli from their *pairwise discrimination probabilities*. The latter are the probabilities with which the judgment 'these two stimuli are different' is chosen over 'these two stimuli are the same.'[2] This application presupposes that a properly defined stimulus space satisfies the fundamental *law of Regular Minimality*. This law allows one to transform (relabel) the stimuli so that the probability ψ**aa** with which **a** is judged to be different from **a** (as explained later, it is better not to say "from itself") is always smaller than ψ**ab** and ψ**ba**, the probabilities with which **a** is judged to be different from any **b** $\neq$ **a**. Aside from the law of Regular Minimality, the application of our DC theory to discrimination probabilities is based on one assumption only: that the differences ψ**ab**$-\psi$**aa** and ψ**ba**$-\psi$**aa**, the so-called *psychometric increments*, are both of them (different) dissimilarity functions.

The application of DC to discrimination probabilities is a radical extension of the theoretical program called Generalized Fechnerian Scaling, which includes as special cases Multidimensional Fechnerian Scaling (Dzhafarov, 2001a-b, 2002a-d, 2003a-b, 2004; Dzhafarov & Colonius, 1999a-b, 2001), Fechnerian Scaling of Continuous Spaces (Dzhafarov & Colonius, 2005a), and Fechnerian Scaling of Discrete Object Sets (Dzhafarov & Colonius, 2005b, 2006b-c). We refer the reader to this literature (especially, Dzhafarov, 2001a, 2002b,d; Dzhafarov & Colonius, 1999a, 2001, 2006c) for historical background and the relation of Generalized Fechnerian Scaling to traditional issues of psychophysics (Fechner's original theory, its experimental and theoretical critiques, "the Fechner problem," geometries of color spaces, Multidimensional Scaling, and others). For recently published related work see Ennis (2006), Iverson (2006), Townsend, Aisbett, Busemeyer, and Assadi (2006), and Zhang (2004, 2006).

In Fechner's original theory every point on a unidimensional stimulus continuum was assigned a measure of its discriminability from its "immediate neighbors," and this discriminability measure, when integrated from one stimulus to another, yielded the subjective distance between them. Since Fechner assumed that the unidimensional continuum of stimuli monotonically corresponded to a unidimensional continuum of "sensations," the discriminability measure in his theory could be derived from greater-less judgments, e.g., by computing the slopes of the probability-of-greater functions at their medians (or, as a crude approximation, the reciprocals of "just-noticeable differences").

In Generalized Fechnerian Scaling the greater-less judgments are replaced with same-different ones, the discriminability measure is computed from the probability-of-different functions, and unidimensional continua are extended into a very broad class of stimulus spaces. This class includes both stimulus spaces in which stimuli can be connected by continuous paths (such as a space of colors) and stimulus spaces comprised

---

[2]These are merely prototypical formulations of the two responses. The notion of discrimination entails a wide variety of judgment types (see Dzhafarov & Colonius, 2006a).

of isolated points (such as a space of color names). Although useful links have been established between the "continuous" and "discrete" computations (Dzhafarov & Colonius, 2005a-b), in the previous versions of Generalized Fechnerian Scaling they remained very different in nature. One would be justified in viewing this as a problem, augmented by difficulties in drawing a clear empirical demarcation line between continuous and discrete stimulus spaces.

The DC theory, by contrast, is applicable to stimulus spaces of entirely arbitrary nature, subject only to the law of Regular Minimality and the above-stipulated assumption about the psychometric increments. To emphasize this fact we call the new, DC-based theory of computing subjective distances from discrimination probabilities *Universal Fechnerian Scaling* (UFS).

Potentially, the idea of DC has an even broader application area. One may reasonably hypothesize that most if not all "dissimilarity-type" measures used in the conventional techniques of Cluster Analysis and Multidimensional Scaling (e.g., an average numerical assessment of dissimilarity) are in fact dissimilarity functions in our special sense.

Whatever the application area, the computation of distances from dissimilarities is a *purely psychological theory* in the technical meaning defined in Dzhafarov and Colonius (2005a): this computation is invariant with respect to all possible relabelings (bijective transformations) of stimuli, thus making no use of their physical properties. Stimuli are entirely and exclusively characterized by their pairwise discrimination probabilities. Thus, the Fechnerian distance between two given stimuli with real-valued representations will remain precisely the same if the unidimensional continuum of stimuli to which they belong is bijectively mapped, say, onto a unit square.

1.1. Notation Conventions

The two abbreviations,

$$\text{DC} = \text{Dissimilarity Cumulation and UFS} = \text{Universal Fechnerian Scaling}$$

are used throughout the paper.

Boldface lowercase letters, $\mathbf{a}$, $\mathbf{b}'$, $\mathbf{x}$, $\mathbf{y}_n$, …, always denote elements of a set of stimuli. Stimuli are merely names (qualitative entities), with no algebraic operations defined on them. If stimuli are represented by real numbers, we use the same symbol for both a stimulus and its numerical representation but use the boldface and lightface type to distinguish them: stimuli $\mathbf{a}$, $\mathbf{x}$, $\mathbf{15}$, ... with values a, x, 15, ….[3]

The set of all stimuli and its subsets are denoted by Gothic letters, $\mathfrak{S}$, $\mathfrak{S}_1^*$, $\mathfrak{s}$, ….

Finite sequences (chains) of stimuli are denoted by uppercase boldface letters, $\mathbf{X}$, $\mathbf{Y}_n$, ….

Real-valued functions of one or more arguments which are elements of a stimulus set are indicated by strings without parentheses: $\psi \mathbf{ab}$, $D\mathbf{abc}$, $D\mathbf{X}_n$, $\Psi^{(\iota)}\mathbf{ab}$, ….

_____

[3] We maintain this rigorous distinction because the DC theory and UFS are being introduced here for the first time. It is generally convenient and innocuous, however, to notationally confuse stimuli with their numerical values.

Infinite sequences $\{x_n\}_{n\in\mathbb{N}}$, $\{\mathbf{x}_n\}_{n\in\mathbb{N}}$, $\{\mathbf{X}_n\}_{n\in\mathbb{N}}$, ... are almost always indicated by their generic elements: sequence x_n, sequence $\mathbf{x}_n$, etc. Convergence of a sequence is always understood as conditioned on $n \to \infty$.

We use the square-bracket notation for intervals of reals (closed, open and half-open): $[a,b]$, $[a,b[$, $]a,b]$, and $]a,b[$. The round brackets, (a,b), always indicate an ordered pair of numbers. The sets of reals, nonnegative reals, and naturals $(1,2,...)$ are denoted traditionally, $\mathbb{R}$, $\mathbb{R}^+$, and $\mathbb{N}$.

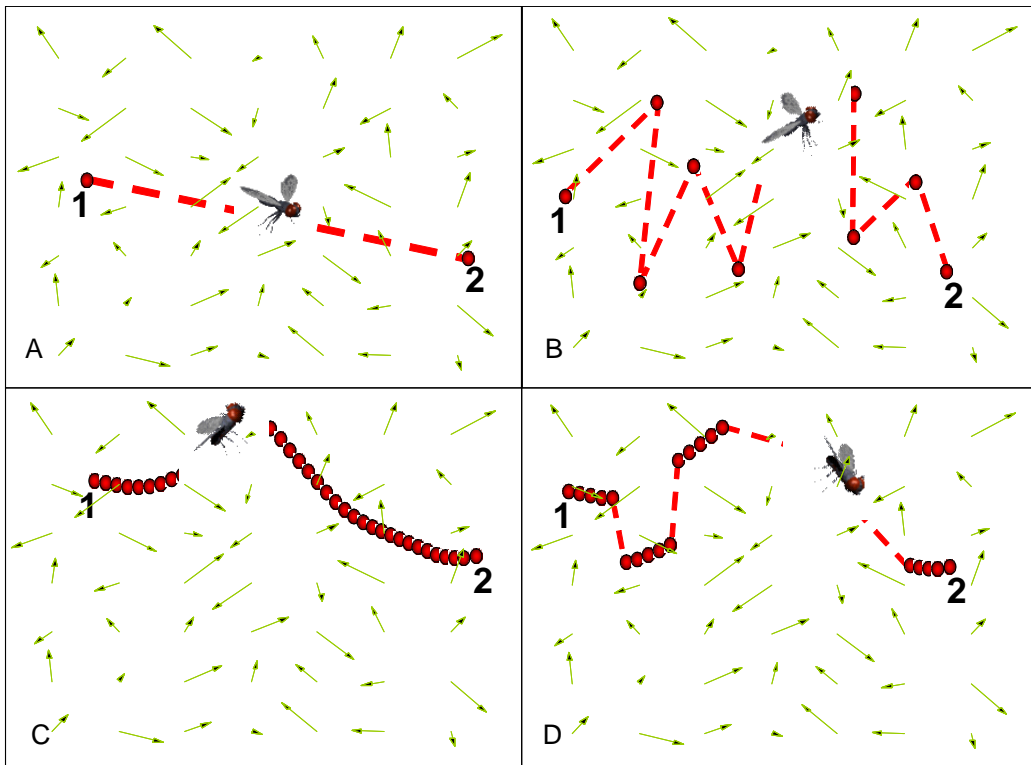## 1.2.   An "Entomological" Metaphor

Figure 1. Different ways of getting from **1** to **2**. (A) Direct flight. (B) Flight with intermediate (instantaneous) stops. (C) "Crawling" (here, infinite chain of infinitesimally small flights). (D) "Crawls" and flights.

Consider a terrain with a complicated but steady pattern of winds blowing over it (Fig. 1). There is a fly that wishes to get from position **1** to position **2**. From the fly's point of view, the pattern of winds is characterized by the amount of time/effort $D\mathbf{ab}$ the fly should expend to fly straight from $\mathbf{a}$ to $\mathbf{b}$, for every pair of positions $(\mathbf{a}, \mathbf{b})$. One can easily see that $D\mathbf{ab}$ thus defined is generally different from $D\mathbf{ba}$, and that, given another position $\mathbf{x}$, $D\mathbf{ab}$ need not be always smaller than $D\mathbf{ax} + D\mathbf{xb}$. Thus, the time/effort of flying from **1** to **2** directly, without intermediate stops (Fig. 1A), may very well be greater than a chain of flights leading from **1** to **2** through intermediate points (Fig. 1B). It is even possible that the fly would do better by "crawling" all the way from **1** to **2** (Fig. 1C) – for the purposes of the present metaphor, "crawling" is understood as a limit case for a chain of flights, as their number increases beyond bounds and

their individual durations (or associated efforts) decrease to zero. It may very well be the case that the shortest (least exerting) path should include a combination of "crawls" and flights, as shown in Fig. 1D.

If the terrain (set of locations) is taken to represent a stimulus set, then the time/effort function $D\mathbf{ab}$ is a metaphor for a dissimilarity function. A stimulus set may or may not be representable as a manifold, may or may not allow for continuous paths (e.g., it can consist of a finite number of isolated points), but the idea of cumulating dissimilarity across all possible finite chains with two given endpoints is applicable universally.
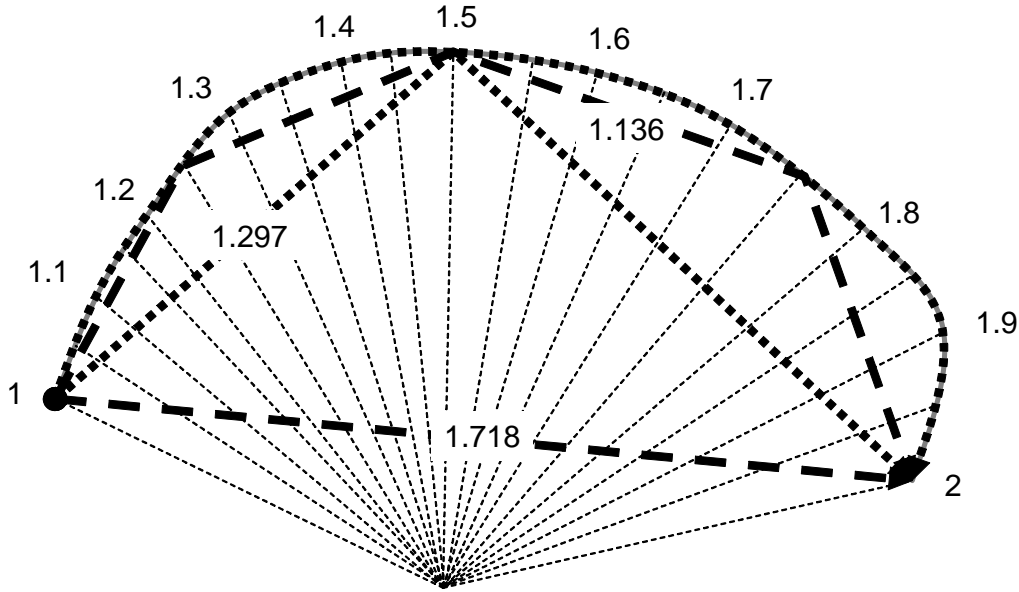

Figure 2. "Lengths" (i.e., cumulated dissimilarities) of different chains of points leading from 1 to 2 in the stimulus space represented by the interval $[1, 2]$. The interval is shown by a curved line to make it graphically distinguishable from the chains of its elements. The dissimilarity function is $D\mathbf{xy} = \left| e^{y-x} - 1 \right|$.

For a numerical example, consider the situation depicted in Fig. 2. The stimulus set here is represented by an interval of reals between 1 and 2, and the dissimilarity function is

$$D\mathbf{xy} = \left| e^{y-x} - 1 \right|,$$

where x, y are numerical values assigned to stimuli $\mathbf{x}, \mathbf{y}$. If the fly flies from $\mathbf{1}$ to $\mathbf{2}$ directly, it takes 1.718 units of time/effort,

$$D\mathbf{12} = \left| e^{2-1} - 1 \right| = 1.718.$$

Flying from $\mathbf{1}$ to $\mathbf{2}$ through point $\mathbf{1.5}$ takes less,

$$D\mathbf{1}\left(\mathbf{1.5}\right)\mathbf{2} = D\mathbf{1}\left(\mathbf{1.5}\right) + D\left(\mathbf{1.5}\right)\mathbf{2} = \left| e^{1.5-1} - 1 \right| + \left| e^{2-1.5} - 1 \right| = 1.297,$$

and the chain $\mathbf{1} - \mathbf{1.25} - \mathbf{1.5} - \mathbf{1.75} - \mathbf{2}$ takes less still,

$$D\mathbf{1}\left(\mathbf{1.25}\right)\left(\mathbf{1.5}\right)\left(\mathbf{1.75}\right)\mathbf{2} = \left| e^{1.25-1} - 1 \right| + ... + \left| e^{2-1.75} - 1 \right| = 1.135.$$

Proceeding in this manner it can be shown (Example 4 in Section 2.5) that the fastest/easiest way to get from **1** to **2** in this case is to "crawl" all the way across the interval: this will take 1 unit of effort,

$$\inf_{k,\mathbf{x}_1...\mathbf{x}_k} D\left(\mathbf{x}_0 = \mathbf{1}\right)\mathbf{x}_1...\mathbf{x}_k\left(\mathbf{x}_{k+1} = \mathbf{2}\right) = \int_1^2 \left.\frac{d\left(e^{y-x}-1\right)}{dy}\right|_{y=x+} dx = \int_1^2 dx = 1.$$

In other words, the infimum of cumulated dissimilarities for all finite chains of stimuli leading from **1** to **2** equals 1. This amount is taken to be the oriented (asymmetric) distance from **1** to **2**:
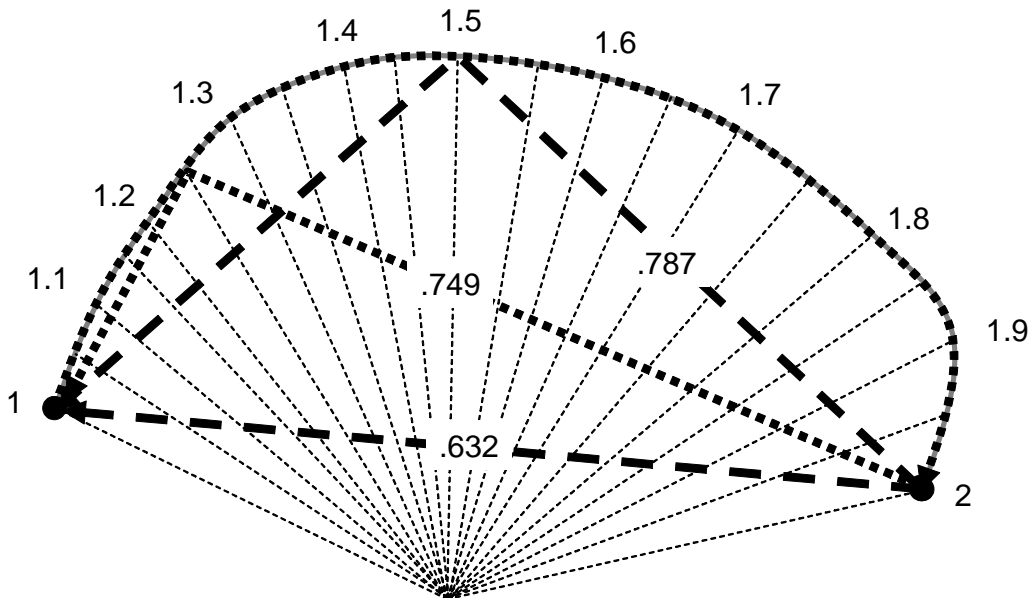
$$G\mathbf{12} = 1.$$

Figure 3. "Lengths" of different chains of points leading from 2 to 1 in the stimulus space of Fig. 2.

The situation is different for getting back, from **2** to **1** (Fig. 3). The direct flight here turns out to be the fastest/easiest path,

$$D\mathbf{21} = \left|e^{1-2}-1\right| = 0.632.$$

Any intermediate point inserted between **2** and **1** only increases the time/effort needed. The infimum of cumulated dissimilarities for all finite chains of stimuli leading from **2** to **1** equals therefore 0.632 units. This amount is taken to be the oriented distance from **2** to **1**:

$$G\mathbf{21} = 0.632.$$

The *overall*, symmetric distance between **1** and **2** can now be computed by adding the two oriented distances "to and from":

$$G\mathbf{12} + G\mathbf{21} = 1.632.$$

A rigorous analysis of this example requires of course that one define the properties of a dissimilarity function (Section 2.4) and prove that $D\mathbf{xy} = |e^{y-x} - 1|$ satisfies these properties on interval $[1, 2]$ (Examples 1 and 4 in Sections 2.1 and 2.5). One should also show that the infimum of dissimilarity values cumulated across all finite chains with fixed endpoints is an oriented distance from the initial to the terminal point (Section 2.5). Finally, one should justify the computation of the overall, symmetric distance by adding together the two oriented distances "to and from," rather than, say, taking their maximum (Sections 1.4 and 2.8). This justification is derived from UFS, by considering the two kinds of psychometric increments as dissimilarity functions, and requiring that the overall (symmetric) distances computed from them coincide (Sections 1.3 and 2.8).

1.3. Psychophysics of Discrimination: Basics

Here, we briefly recapitulate some of the basic concepts and assumptions underlying the theory of same-different discrimination probabilities. The discussion is illustrated by Fig. 4. A detailed description and examples can be found in Dzhafarov (2002d, 2003a) and Dzhafarov and Colonius (2005a, 2006a).

1.3.1. Two observation areas

(See matrix ψ^* in Fig. 4.) The arguments $\mathbf{x}$ and $\mathbf{y}$ of the discrimination probability function

$$\psi^*\mathbf{xy} = \Pr\left[\mathbf{x} \text{ and } \mathbf{y} \text{ are judged to be different}\right] \tag{1}$$

belong to two distinct *observation areas*,

$$\psi^* : \mathfrak{S}_1^* \times \mathfrak{S}_2^* \mapsto [0, 1]. \tag{2}$$

Thus, $\mathfrak{S}_1^*$ (the *first observation area*) may represent stimuli presented chronologically first or on the left, whereas $\mathfrak{S}_2^*$ (the *second observation area*) designates stimuli presented, respectively, chronologically second or on the right. The adjectives "first" and "second" refer to the ordinal positions of stimulus symbols within a pair $(\mathbf{x}, \mathbf{y})$.

1.3.2. Psychological equality and reduced observation areas

(See the transition from matrix ψ^* to matrix $\tilde{\psi}$ in Fig. 4.) For $\mathbf{x}, \mathbf{x}' \in \mathfrak{S}_1^*$, we say that the two stimuli are *psychologically equal* (or metameric) if $\psi^*\mathbf{xy} = \psi^*\mathbf{x}'\mathbf{y}$ for any $\mathbf{y} \in \mathfrak{S}_2^*$. Analogously, the psychological equality for $\mathbf{y}, \mathbf{y}' \in \mathfrak{S}_2^*$ is defined by $\psi^*\mathbf{xy} = \psi^*\mathbf{xy}'$, for any $\mathbf{x} \in \mathfrak{S}_1^*$. The *reduced first observation area* $\mathfrak{S}_1$ is the factor set of $\mathfrak{S}_1^*$ with respect to the psychological equality on $\mathfrak{S}_1^*$; and analogously for the *reduced second observation area* $\mathfrak{S}_2$. Put differently, the reduction of the observation areas consists in relabeling their elements so that psychologically equal stimuli receive identical labels and are no longer distinguished. The discrimination probability function ψ^* can then be redefined as

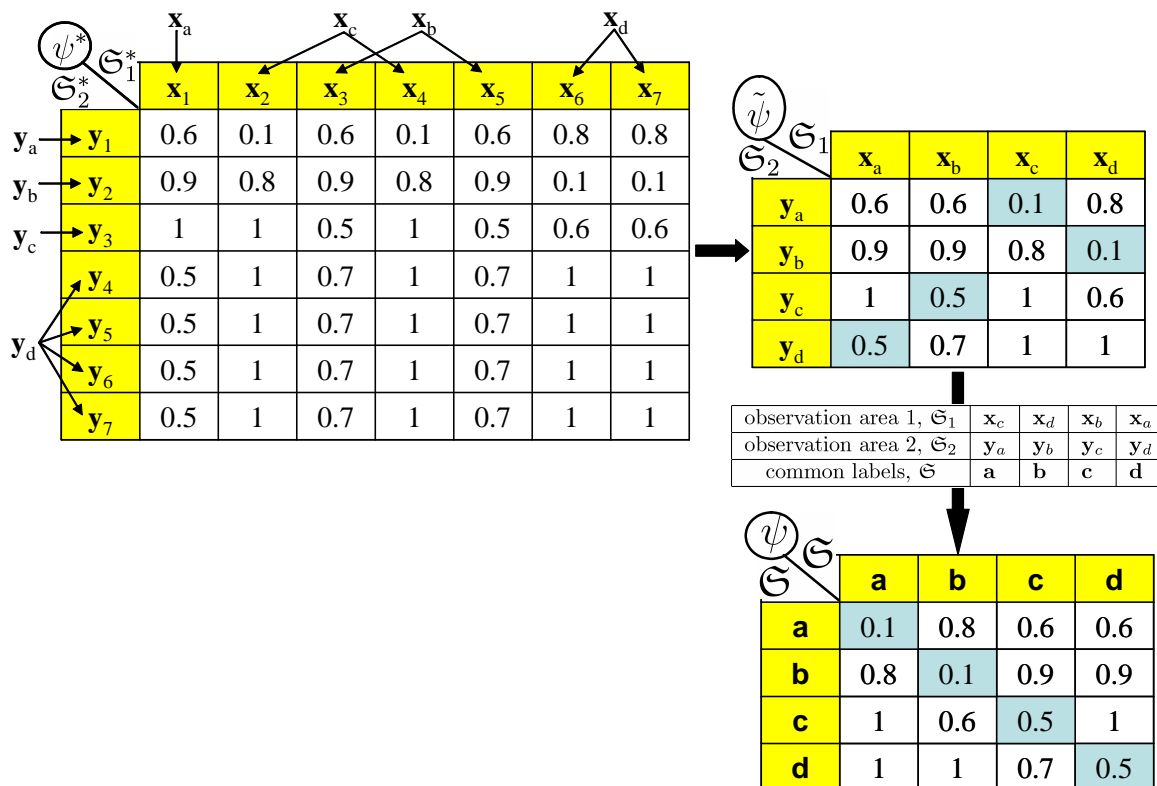$$\tilde{\psi} : \mathfrak{S}_1 \times \mathfrak{S}_2 \mapsto [0, 1]. \tag{3}$$

$\psi^* \; \mathfrak{S}_1^*$ / $\mathfrak{S}_2^*$	$\mathbf{x}_1$	$\mathbf{x}_2$	$\mathbf{x}_3$	$\mathbf{x}_4$	$\mathbf{x}_5$	$\mathbf{x}_6$	$\mathbf{x}_7$
$\mathbf{y}_1$	0.6	0.1	0.6	0.1	0.6	0.8	0.8
$\mathbf{y}_2$	0.9	0.8	0.9	0.8	0.9	0.1	0.1
$\mathbf{y}_3$	1	1	0.5	1	0.5	0.6	0.6
$\mathbf{y}_4$	0.5	1	0.7	1	0.7	1	1
$\mathbf{y}_5$	0.5	1	0.7	1	0.7	1	1
$\mathbf{y}_6$	0.5	1	0.7	1	0.7	1	1
$\mathbf{y}_7$	0.5	1	0.7	1	0.7	1	1

$\tilde{\psi} \; \mathfrak{S}_1$ / $\mathfrak{S}_2$	$\mathbf{x}_a$	$\mathbf{x}_b$	$\mathbf{x}_c$	$\mathbf{x}_d$
$\mathbf{y}_a$	0.6	0.6	0.1	0.8
$\mathbf{y}_b$	0.9	0.9	0.8	0.1
$\mathbf{y}_c$	1	0.5	1	0.6
$\mathbf{y}_d$	0.5	0.7	1	1

observation area 1, $\mathfrak{S}_1$	$\mathbf{x}_c$	$\mathbf{x}_d$	$\mathbf{x}_b$	$\mathbf{x}_a$
observation area 2, $\mathfrak{S}_2$	$\mathbf{y}_a$	$\mathbf{y}_b$	$\mathbf{y}_c$	$\mathbf{y}_d$
common labels, $\mathfrak{S}$	a	b	c	d

$\psi \; \mathfrak{S}$ / $\mathfrak{S}$	a	b	c	d
a	0.1	0.8	0.6	0.6
b	0.8	0.1	0.9	0.9
c	1	0.6	0.5	1
d	1	1	0.7	0.5

Figure 4. A toy example adopted from Dzhafarov & Colonius (2006a). The transformation from $\left(\mathfrak{S}_1^*, \mathfrak{S}_2^*, \psi^*\right)$ to $\left(\mathfrak{S}_1, \mathfrak{S}_2, \tilde{\psi}\right)$ is the result of "lumping together" psychologically equal stimuli (e.g., the stimuli $\mathbf{y}_4, \mathbf{y}_5, \mathbf{y}_6, \mathbf{y}_7$ are psychologically equal in $\mathfrak{S}_2^*$, stimuli $\mathbf{x}_2$ and $\mathbf{x}_4$ are psychologically equal in $\mathfrak{S}_1^*$). The space $\left(\mathfrak{S}_1, \mathfrak{S}_2, \tilde{\psi}\right)$ satisfies the Regular Minimality condition (the minimum in each row is also the minimum in its column) because of which $\left(\mathfrak{S}_1, \mathfrak{S}_2, \tilde{\psi}\right)$ can be canonically transformed into $(\mathfrak{S}, \psi)$, by means of the transformation table shown in between.

The adjective "reduced" is usually dropped, and $\mathfrak{S}_1, \mathfrak{S}_2$ are referred to simply as the first and second observation areas. The elements of $\mathfrak{S}_1, \mathfrak{S}_2$ are referred to as stimuli even through they are in fact equivalence classes of the elements of $\mathfrak{S}_1^*, \mathfrak{S}_2^*$, respectively.

1.3.3. The Law of Regular Minimality

(See matrix $\tilde{\psi}$ in Fig. 4.) We assume that there are functions $\mathbf{h}: \mathfrak{S}_1 \mapsto \mathfrak{S}_2$ and $\mathbf{g}: \mathfrak{S}_2 \mapsto \mathfrak{S}_1$ such that

$$
\begin{array}{ll}
(\mathcal{P}_1) & \tilde{\psi}\mathbf{x}\left[\mathbf{h}\left(\mathbf{x}\right)\right] < \tilde{\psi}\mathbf{x}\mathbf{y} \text{ for all } \mathbf{y} \neq \mathbf{h}\left(\mathbf{x}\right) \\
(\mathcal{P}_2) & \tilde{\psi}\left[\mathbf{g}\left(\mathbf{y}\right)\right]\mathbf{y} < \tilde{\psi}\mathbf{x}\mathbf{y} \text{ for all } \mathbf{x} \neq \mathbf{g}\left(\mathbf{y}\right) \\
(\mathcal{P}_3) & \mathbf{h} \equiv \mathbf{g}^{-1}
\end{array}
\tag{4}
$$

Clearly, this implies that $\mathbf{h}$ and $\mathbf{g}$ are bijections. Stimulus $\mathbf{y} = \mathbf{h}\left(\mathbf{x}\right) \in \mathfrak{S}_2$ is called the *Point of Subjective Equality* (PSE) for $\mathbf{x} \in \mathfrak{S}_1$; analogously, $\mathbf{x} = \mathbf{g}\left(\mathbf{y}\right) \in \mathfrak{S}_1$ is the PSE for $\mathbf{y} \in \mathfrak{S}_2$. The law of Regular Minimality states therefore that every stimulus in each of the observation areas has a unique PSE in the other observation area, and that $\mathbf{y}$ is the PSE for $\mathbf{x}$ if and only if $\mathbf{x}$ is the PSE for $\mathbf{y}$. In some contexts the law of Regular Minimality is an empirical assumption, but it can also serve as a criterion for a properly defined stimulus space. For a detailed discussion of the law and its critiques see Dzhafarov (2002d, 2003a, 2006), Dzhafarov and Colonius (2006a), and Ennis (2006).

1.3.4. Canonical transformation and psychometric increments

(See the transition from matrix $\tilde{\psi}$ to matrix ψ in Fig. 4.) Due to the law of Regular Minimality, one can always relabel the stimuli in $\mathfrak{S}_1$ and/or $\mathfrak{S}_2$ so that any two mutual PSEs receive one and the same label. In other words, one can always bijectively map $\mathfrak{S}_1 \mapsto \mathfrak{S}$ and $\mathfrak{S}_2 \mapsto \mathfrak{S}$ so that $\mathbf{x} \mapsto \mathbf{a}$ and $\mathbf{y} \mapsto \mathbf{a}$ if and only if $\mathbf{x} \in \mathfrak{S}_1$ and $\mathbf{y} \in \mathfrak{S}_2$ are mutual PSEs: $\mathbf{y} = \mathbf{h}\left(\mathbf{x}\right), \mathbf{x} = \mathbf{g}\left(\mathbf{y}\right)$. The set of labels $\mathfrak{S}$ is called a *canonically transformed* stimulus set. Its elements too, for simplicity, are referred to as stimuli. The discrimination probability function $\tilde{\psi}$ can now be presented in a *canonical form*,

$$
\psi: \mathfrak{S} \times \mathfrak{S} \mapsto [0, 1],
\tag{5}
$$

with the property

$$
\psi\mathbf{a}\mathbf{a} < \min\left\{\psi\mathbf{a}\mathbf{b}, \psi\mathbf{b}\mathbf{a}\right\}
\tag{6}
$$

for any $\mathbf{a}$ and $\mathbf{b} \neq \mathbf{a}$. Note that the first and the second $\mathbf{a}$ in $\psi\mathbf{a}\mathbf{a}$ may very well refer to physically different stimuli (equivalence classes of stimuli): hence one should exercise caution in referring to $\psi\mathbf{a}\mathbf{a}$ as the probability with which $\mathbf{a}$ is discriminated from "itself."

Not to have the presentation too schematic, Fig. 5 illustrates the notion of a canonical transformation on a real data matrix.

For the canonically transformed function ψ, the *psychometric increments of the first and second kind* are defined as, respectively,

$$
\Psi^{(1)}\mathbf{a}\mathbf{b} = \psi\mathbf{a}\mathbf{b} - \psi\mathbf{a}\mathbf{a}
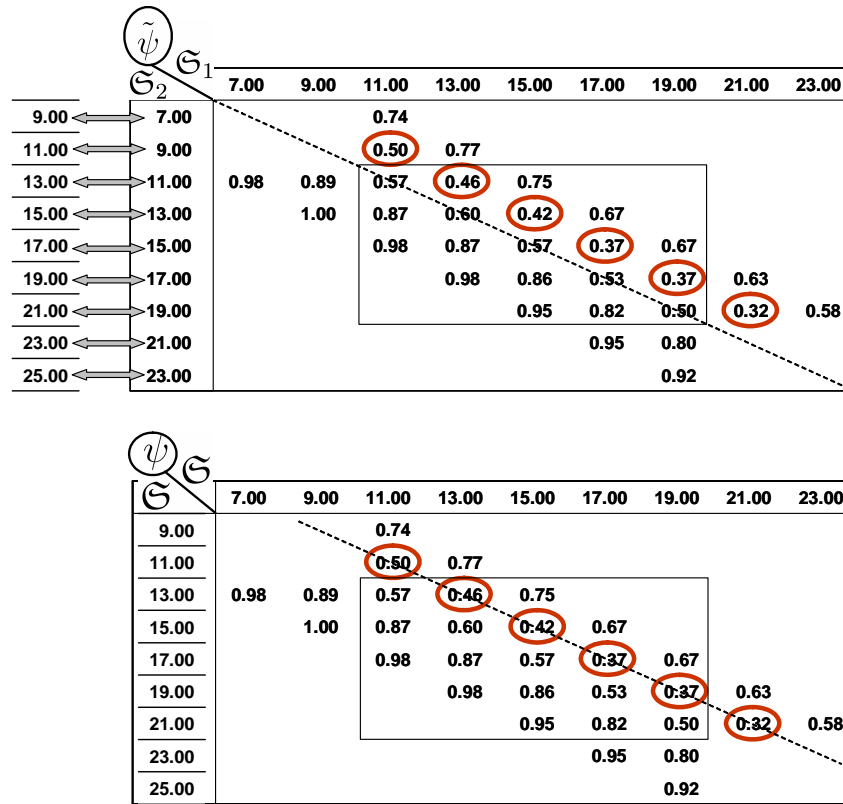\tag{7}
$$

Figure 5. An illustration for a canonical transformation on a data matrix for Experiment A described in Dzhafarov & Colonius (2005a). The stimuli are horizontal line segments presented on the left ($\mathfrak{S}_1 = \mathfrak{S}_1^*$) and right ($\mathfrak{S}_2 = \mathfrak{S}_2^*$), their lengths in pixels shown on the margins (1 px $\approx$ 0.86 min arc). The probability estimates are shown for one participant, they are based on approximately 550 replications per pair in the central area and 275 replications per pair outside it. Within the precision of the experiment Regular Minimality is satisfied: the minima in each row (encircled) are also the minima in their columns. The PSE relation here is $y = x - 2$. A canonical transformation therefore can be effected by replacing each y-value in $\mathfrak{S}_2$ by $y + 2$, as shown (more generally, by any bijective transformation $x \mapsto z$, $y \mapsto z + 2$). The matrix below is in a canonical form: the PSE pairs lie all on the main diagonal. Note the Nonconstant Self-Dissimilarity property (the diagonal values are clearly different) and the lack of symmetry with respect to the redefined main diagonal.

and

$$\Psi^{(2)}\mathbf{ab} = \psi\mathbf{ba} - \psi\mathbf{aa}. \tag{8}$$

Due to the canonical form of ψ these quantities are always positive for $\mathbf{b} \neq \mathbf{a}$.

1.4. The Logic of Fechnerian Scaling

1.4.1. Stimulus metric is computed within rather than across observation areas

A canonical transformation of stimulus space is more than a notational convenience. The identification of the two observation areas with one and the same set $\mathfrak{S}$ is the only way to speak of a stimulus metric: a metric cannot be defined on a Cartesian product of two distinct sets. In Fechnerian Scaling the overall (symmetric) Fechnerian distance $G^*\mathbf{ab}$ between two points in the canonical space $\mathfrak{S}$ is interpreted as the subjective distance between the corresponding two stimuli *within* each of the observation areas (Dzhafarov & Colonius, 2005a, 2006b). That is, $G^*\mathbf{ab}$ is the distance between any two physical stimuli mapped into labels $\mathbf{a}$ and $\mathbf{b}$ from the first observation area, $\mathfrak{S}_1^*$ or $\mathfrak{S}_1$ (say, both presented chronologically first or on the left). And of course, $G^*\mathbf{ab}$ is also interpretable as the distance between any two stimuli mapped into labels $\mathbf{a}$ and $\mathbf{b}$ from the second observation area, $\mathfrak{S}_2^*$ or $\mathfrak{S}_2$ (both presented chronologically second or on the right). This interpretation clarifies, in particular, why it is both meaningful and necessary to have $G^*\mathbf{aa} = 0$: the two $\mathbf{a}$'s here refer to one and the same stimulus (say, a given color presented in a given way) or two psychologically equal stimuli (two metameric colors presented in one and the same way). By contrast, the discrimination probabilities $\psi\mathbf{ab}$ are always defined *across* two observation areas, and $\psi\mathbf{aa}$ refers to the probability with which a stimulus mapped into label $\mathbf{a}$ from the first observation area is discriminated from a stimulus mapped into label $\mathbf{a}$ from the second observation area (which probability therefore need not be zero and generally varies from one label to another).[4]

L	a	b	c	d
a	a	acba	aca	ada
b	bacb	b	bcb	bdcb
c	cac	cbc	c	cdc
d	dad	dcbd	dcd	d

G*	a	b	c	d
a	0	1.3	1	1
b	1.3	0	0.9	1.1
c	1	0.9	0	0.7
d	1	1.1	0.7	0

Figure 6. Fechnerian distances (matrix G^*) and geodesic loops (matrix L) computed from matrix ψ in Fig. 4 (adopted from Dzhafarov & Colonius, 2006b,c, where the computational details are given). The geodesic loops are the shortest chains "to and from": e.g., the shortest way of getting from $\mathbf{d}$ to $\mathbf{b}$ is $\mathbf{dcb}$, and the shortest way back is $\mathbf{bd}$ ("direct flight"), comprising together the loop $\mathbf{dcbd}$ of length 1.1.

---

[4] The non-invariance of $\psi\mathbf{aa}$ with respect to $\mathbf{a}$, referred to as the Nonconstant Self-Dissimilarity property, when combined with the law of Regular Minimality plays an important role in the analysis of several schemes of perceptual analysis (Dzhafarov, 2002b, 2003a-b, 2006; Dzhafarov & Colonius, 2006a).

Referring to Fig. 6, to say that the (overall, symmetric) distance between **d** and **b** equals 1.1 is equivalent to saying (with reference to the correspondence table in Fig. 4) that 1.1 is the distance between $\mathbf{x}_a$ and $\mathbf{x}_d$, both of which are within the (reduced) observation area $\mathfrak{S}_1$; and also that 1.1 is the distance between $\mathbf{y}_d$ and $\mathbf{y}_b$, both of them in the (reduced) observation area $\mathfrak{S}_2$. These statements in their turn are equivalent to saying (with reference to matrix ψ^* in Fig. 4) that 1.1 is the distance between $\mathbf{x}_1$ and either of the stimuli $\{\mathbf{x}_6, \mathbf{x}_7\}$, all in the original observation area $\mathfrak{S}_1^*$; and also that 1.1 is the distance between $\mathbf{y}_2$ and any of the stimuli $\{\mathbf{y}_4, \mathbf{y}_5, \mathbf{y}_6, \mathbf{y}_7\}$, all in the original observation area $\mathfrak{S}_2^*$.

1.4.2. Stimulus metric is symmetric rather than oriented

The within-observation-area interpretation for Fechnerian distances makes it clear that a meaningful Fechnerian distance must be symmetric rather than oriented: the order of the two stimuli belonging to one and the same observation area has no operational meaning. Thus, stimuli $\mathbf{y}_7$ and $\mathbf{y}_2$ in the original observation area $\mathfrak{S}_2^*$ (say, both presented chronologically second or both presented on the right) can never be "presented together" or "directly compared": the distance between them is being computed based on the difference between the discrimination probability functions $\psi^*\mathbf{x}\mathbf{y}_7$ and $\psi^*\mathbf{x}\mathbf{y}_2$, when $\mathbf{y}_7$ and $\mathbf{y}_2$ are compared with all possible stimuli in the other observation area (all chronologically first stimuli, or all stimuli presented on the left). Clearly, the ordered pairs $(\mathbf{y}_7, \mathbf{y}_2)$ and $(\mathbf{y}_2, \mathbf{y}_7)$ are operationally indistinguishable. The logic of Fechnerian Scaling requires therefore that we compute a *single symmetric distance function*.[5] As we know, this is being done by arithmetically adding the oriented distances "to" and "from," computed by means of what we call the standard procedure of the DC theory (Sections 1 and 1.2). The rationale for this particular symmetrization scheme is as follows.

1.4.3. Stimulus metric is the same for the two kinds of psychometric increments

The main assumption of UFS about the psychometric increments of a canonically transformed ψ is that both of them are dissimilarity functions (as defined in Sections 2.1-2.4). One can then compute the oriented distance $G_1\mathbf{ab}$ from the psychometric increments of the first kind, $\Psi^{(1)}$, by considering all possible finite chains of stimuli $\mathbf{x}_1...\mathbf{x}_k$ for all possible k and putting

$$G_1\mathbf{ab} = \inf_{k, \mathbf{x}_1...\mathbf{x}_k} \left[\Psi^{(1)}\mathbf{ax}_1 + \Psi^{(1)}\mathbf{x}_1\mathbf{x}_2 + ... + \Psi^{(1)}\mathbf{x}_k\mathbf{b} \right]. \tag{9}$$

The overall Fechnerian distance is then computed as

$$G^*\mathbf{ab} = G_1\mathbf{ab} + G_1\mathbf{ba}. \tag{10}$$

It is clear from symmetry considerations, however, that we can equally well use $\Psi^{(2)}$ in place of $\Psi^{(1)}$, computing first

$$G_2\mathbf{ab} = \inf_{k, \mathbf{x}_1...\mathbf{x}_k} \left[\Psi^{(2)}\mathbf{ax}_1 + \Psi^{(2)}\mathbf{x}_1\mathbf{x}_2 + ... + \Psi^{(2)}\mathbf{x}_k\mathbf{b} \right] \tag{11}$$

---

[5] We are by no means suggesting that the oriented distances allow for no meaningful interpretation, only that they are not meaningful distances *within* observation areas.

and then adding together

$$G_2 \mathbf{ab} + G_2 \mathbf{ba} = G^* \mathbf{ab}. \tag{12}$$

Theorem 14 in Section 2.8 (essentially replicating a proof presented in Dzhafarov & Colonius, 2006b-c) shows that the resulting quantity, $G^* \mathbf{ab}$, is always the same for the two computations, $\Psi^{(1)}$-based and $\Psi^{(2)}$-based.[6]

For a numerical example, take again points $\mathbf{d}$ and $\mathbf{b}$ in Fig. 4 (matrix ψ) and Fig. 6. $G_1 \mathbf{db}$ can be shown (by trying all possible chains leading from $\mathbf{d}$ to $\mathbf{b}$) to be the $\Psi^{(1)}$-length of the chain $\mathbf{dcb}$:

$$\Psi^{(1)} \mathbf{dc} + \Psi^{(1)} \mathbf{cb} = 0.3.$$

For the reverse direction, $G_1 \mathbf{bd}$ is found to equal the $\Psi^{(1)}$-length of the chain $\mathbf{bd}$ ("direct flight"):

$$\Psi^{(1)} \mathbf{bd} = 0.8.$$

The sum of these two quantities gives us the value 1.1 we find in the cells $\mathbf{db}$ and $\mathbf{bd}$ of the matrix G^*. Analogously,

$$G_2 \mathbf{db} = \Psi^{(2)} \mathbf{db} = 0.4,$$

while

$$G_2 \mathbf{bd} = \Psi^{(2)} \mathbf{bc} + \Psi^{(2)} \mathbf{cd} = 0.7,$$

yielding the same sum of 1.1. The overall distances are equal even though the oriented distances $G_1 \mathbf{db}$, $G_1 \mathbf{bd}$, $G_2 \mathbf{db}$, and $G_2 \mathbf{bd}$ are all distinct.

1.4.4. Addition is (essentially) the only universally applicable symmetrization scheme

If one imposes no a priori restrictions on possible stimulus spaces, the addition of oriented distances

$$G^* \mathbf{ab} = G_1 \mathbf{ab} + G_1 \mathbf{ba} = G_2 \mathbf{ab} + G_2 \mathbf{ba} \tag{13}$$

is in fact the only reasonable way of computing one and the same symmetric metric from the pairs $(G_1 \mathbf{ab}, G_1 \mathbf{ba})$ and $(G_2 \mathbf{ab}, G_2 \mathbf{ba})$. To demonstrate this, consider a general symmetrization scheme, that is, a function $f(x, y)$ satisfying the conditions

$$f(x, y) = f(y, x) \tag{14}$$

and

$$f(G_1 \mathbf{ab}, G_1 \mathbf{ba}) = f(G_2 \mathbf{ab}, G_2 \mathbf{ba}).^7 \tag{15}$$

---

[6] In previous publications we interpreted the $\Psi^{(1)}$-based distances as those among stimuli within the first observation area and $\Psi^{(2)}$-based distances as those among stimuli within the second observation area. These attributions, however, are arbitrary and could as well be reversed.

[7] The argument to follow in fact makes no use of the symmetry condition, and it works for a more general form of the second condition: $f_1(G_1 \mathbf{ab}, G_1 \mathbf{ba}) = f_2(G_2 \mathbf{ab}, G_2 \mathbf{ba})$. Distinguishing f_1 and f_2, however, is not meaningful, because either of the observation areas can be labelled first or second arbitrarily.

Consider a two-element stimulus set $\{\mathbf{a}, \mathbf{b}\}$ endowed with four probabilities $\psi\mathbf{aa}$, $\psi\mathbf{ab}$, $\psi\mathbf{ba}$, $\psi\mathbf{bb}$. Let these probabilities be allowed to attain all possible values, subject only to the Regular Minimality constraint,

$$\max\{\psi\mathbf{aa}, \psi\mathbf{bb}\} < \min\{\psi\mathbf{ab}, \psi\mathbf{ba}\}.$$

The oriented distances in this space are

$$G_1\mathbf{ab} = \Psi^{(1)}\mathbf{ab} = \psi\mathbf{ab} - \psi\mathbf{aa}$$

and

$$G_2\mathbf{ab} = \Psi^{(2)}\mathbf{ab} = \psi\mathbf{ba} - \psi\mathbf{aa}.$$

Observe that for a given value of $G^*\mathbf{ab} = s \in]0, 2]$, each of the oriented distances takes values within interval

$$\mathbb{D}_s = \begin{cases}]0, s[& \text{if } s \leq 1 \\ ]s - 1, 1] & \text{if } s > 1 \end{cases}$$

It is easy to see that for any s and any $z \in \mathbb{D}_s$ one can find probabilities $\psi\mathbf{aa}$, $\psi\mathbf{ab}$, $\psi\mathbf{ba}$, $\psi\mathbf{bb}$ (subject to Regular Minimality) satisfying

$$
\begin{aligned}
G_1\mathbf{ab} &= \psi\mathbf{ab} - \psi\mathbf{aa} = s/2 \\
G_1\mathbf{ba} &= \psi\mathbf{ba} - \psi\mathbf{bb} = s/2 \\
G_2\mathbf{ab} &= \psi\mathbf{ba} - \psi\mathbf{aa} = s - z \\
G_2\mathbf{ba} &= \psi\mathbf{ab} - \psi\mathbf{bb} = z.
\end{aligned}
$$

(For instance, if $z \geq s/2$, one can always put $\psi\mathbf{bb} = 0$, $\psi\mathbf{aa} = z - s/2$, $\psi\mathbf{ba} = s/2$, $\psi\mathbf{ab} = z$.) It follows that

$$f(s/2, s/2) = f(z, s - z)$$

for all $z \in \mathbb{D}_s$, which means that f only depends on s and not on z,

$$f(z, s - z) = g(s).$$

In principle g does not have to be an identity function, it can be any function with the following properties:

$$
\begin{aligned}
g(0) &= 0 \\
a + b &\geq c \implies g(a) + g(b) \geq g(c).
\end{aligned}
$$

We can call such functions "metric-preserving," because $g(G^*\mathbf{ab})$ is a metric whenever $G^*\mathbf{ab}$ is. For instance, $[G^*\mathbf{ab}]^p$ is always "metric-preserving" for $0 < p \leq 1$. We will not need to get into a theory of such transformations, however, once we adopt the following natural requirement: if $G_1\mathbf{ab} = G_1\mathbf{ba}$ for all $\mathbf{a}, \mathbf{b}$ (i.e., the oriented distance is already symmetric) then the symmetric metric $g(G^*\mathbf{ab})$ has to be a multiple of $G_1\mathbf{ab}$ (and analogously for G_2). With this requirement adopted,

$$g(G^*\mathbf{ab}) = cG^*\mathbf{ab}, c > 0.$$

It makes little difference which value of c to use. One might think $c = \frac{1}{2}$ to be a convenient choice, but we prefer $c = 1$, as this choice affords the following attractive interpretation: $G^*\mathbf{ab}$ is the infimum of $\Psi^{(1)}$-lengths (or $\Psi^{(2)}$-lengths, the two computations yield the same result) of all finite *closed chains* that contain points $\mathbf{a}$ and $\mathbf{b}$ (see Fig. 7 and matrix L in Fig. 6). That is,

$$G^*\mathbf{ab} = \inf_{\substack{k,\mathbf{x}_1...\mathbf{x}_k \\ l,\mathbf{y}_1...\mathbf{y}_l}} \left[\Psi^{(1)}\mathbf{ax}_1 + \Psi^{(1)}\mathbf{x}_1\mathbf{x}_2 + ... + \Psi^{(1)}\mathbf{x}_k\mathbf{b} + \Psi^{(1)}\mathbf{by}_1 + \Psi^{(1)}\mathbf{y}_1\mathbf{y}_2 + ... + \Psi^{(1)}\mathbf{y}_l\mathbf{a} \right] \tag{16}$$
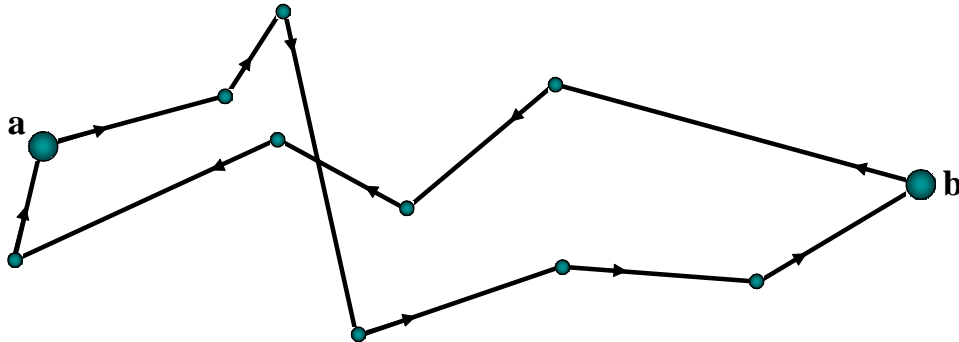
Figure 7. A finite closed loop containing $\mathbf{a}$ and $\mathbf{b}$. The infimum of $\Psi^{(\iota)}$-lengths ($\iota = 1$ or 2) of all such loops is the overall (symmetric) Fechnerian distance between $\mathbf{a}$ and $\mathbf{b}$.

2.   Systematic Development

The DC theory can now be presented systematically. We introduce the notion of dissimilarity in two steps. First we stipulate three conditions that define the notion of a *uniform deviation function*. This notion is sufficient to impose a *topology* and *uniformity* on stimulus space. Then we introduce an additional property that defines a *dissimilarity function*. At the end of the development we return to our main psychological application (discrimination probabilities), with psychometric increments of the two kinds playing the role of two dissimilarity functions.

Although in the application to discrimination probabilities the codomain of a dissimilarity function is $[0, 1]$, we also have to keep in mind the possible non-probabilistic applications mentioned in Section 1, as well as the possibility of using nonlinearly transformed probabilities in Fechnerian Scaling, as discussed in Dzhafarov and Colonius (2005a-c, 2006b). In the general DC theory therefore the codomain of a dissimilarity function is not restricted (within $\mathbb{R}^+$).

2.1.   Uniform Deviation Function

DEFINITION 1 Function $D : \mathfrak{S} \times \mathfrak{S} \mapsto \mathbb{R}$ is a (uniform) deviation function if it has the following properties:

$\mathcal{D}1.$ $\mathbf{a} \neq \mathbf{b} \Longrightarrow D\mathbf{ab} > 0;$

$\mathcal{D}$2. $D\mathbf{aa} = 0$;

$\mathcal{D}$3. (Uniform Continuity) If $D\mathbf{a}_n\mathbf{a}'_n \to 0$ and $D\mathbf{b}_n\mathbf{b}'_n \to 0$, then $D\mathbf{a}'_n\mathbf{b}'_n - D\mathbf{a}_n\mathbf{b}_n \to 0$.
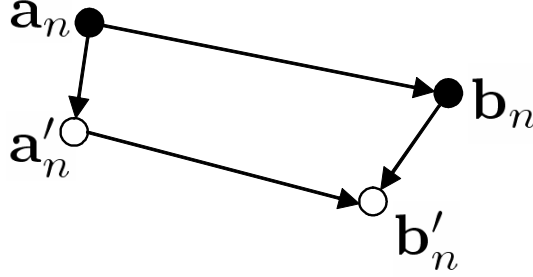

Figure 8. An illustration for Property $\mathcal{D}$3. Consider an infinite sequence of quadrilaterals $\mathbf{a}_1\mathbf{a}'_1\mathbf{b}'_1\mathbf{b}_1$, $\mathbf{a}_2\mathbf{a}'_2\mathbf{b}'_2\mathbf{b}_2$, ... arbitrarily situated with respect to each other but such that the D-lengths of the sides $\mathbf{a}_n\mathbf{a}'_n$ and $\mathbf{b}_n\mathbf{b}'_n$ (oriented as shown by the arrows) gradually vanish. Then the D-lengths of the sides $\mathbf{a}_n\mathbf{b}_n$ and $\mathbf{a}'_n\mathbf{b}'_n$ (again, in the direction of the arrows) gradually converge to each other.

(See Fig. 8.) An equivalent way of formulating the uniform continuity of D is

$$\lim_{\substack{D\mathbf{aa}'\to 0 \\ D\mathbf{bb}'\to 0}} \left(D\mathbf{a}'\mathbf{b}' - D\mathbf{ab}\right) = 0, \tag{17}$$

or, in extenso, for every $\varepsilon > 0$ one can find a $\delta > 0$ such that $|D\mathbf{a}'\mathbf{b}' - D\mathbf{ab}| < \varepsilon$ whenever $D\mathbf{aa}' < \delta$ and $D\mathbf{bb}' < \delta$.[8]

If D is a (symmetric) metric, then it is a deviation function, with the uniform continuity property holding as a theorem. Indeed, by the triangle inequality,

$$D\mathbf{a}_n\mathbf{a}'_n + D\mathbf{a}'_n\mathbf{b}'_n + D\mathbf{b}'_n\mathbf{b}_n \;\geq\; D\mathbf{a}_n\mathbf{b}_n,$$
$$D\mathbf{a}'_n\mathbf{a}_n + D\mathbf{a}_n\mathbf{b}_n + D\mathbf{b}_n\mathbf{b}'_n \;\geq\; D\mathbf{a}'_n\mathbf{b}'_n,$$

whence

$$D\mathbf{a}_n\mathbf{a}'_n + D\mathbf{b}'_n\mathbf{b}_n \;\geq\; D\mathbf{a}_n\mathbf{b}_n - D\mathbf{a}'_n\mathbf{b}'_n$$
$$D\mathbf{a}'_n\mathbf{a}_n + D\mathbf{b}_n\mathbf{b}'_n \;\geq\; D\mathbf{a}'_n\mathbf{b}'_n - D\mathbf{a}_n\mathbf{b}_n.$$

Due to the symmetry condition, the two inequalities can be combined as

$$|D\mathbf{a}'_n\mathbf{b}'_n - D\mathbf{a}_n\mathbf{b}_n| \leq D\mathbf{a}_n\mathbf{a}'_n + D\mathbf{b}_n\mathbf{b}'_n,$$

---

[8]That $\mathcal{D}$3 follows from this statement is obvious. To see the reverse implication, assume there is an $\varepsilon > 0$ such that for any $\delta > 0$ one can find $\mathbf{a}_\delta, \mathbf{a}'_\delta, \mathbf{b}_\delta, \mathbf{b}'_\delta$ for which $D\mathbf{a}_\delta\mathbf{a}'_\delta < \delta$ and $D\mathbf{b}_\delta\mathbf{b}'_\delta < \delta$ but $|D\mathbf{a}'_\delta\mathbf{b}'_\delta - D\mathbf{a}_\delta\mathbf{b}_\delta| \geq \varepsilon$. Then choosing any sequence $\delta_n \to 0$ and denoting $\mathbf{a}_{\delta_n} = \mathbf{a}_n$, $\mathbf{a}'_{\delta_n} = \mathbf{a}'_n$, etc., one would arrive at a contradiction with $\mathcal{D}$3. We often formulate limit propositions in terms of sequences and reformulate them in ε-δ terms when convenient, omitting a proof of equivalence which is always analogous to the one just given.

and $\mathcal{D}3$ follows. It is easy to see that if D is an oriented metric, then it is a deviation function if and only if it additionally satisfies the condition

$$D\mathbf{a}_n\mathbf{a}'_n \to 0 \implies D\mathbf{a}'_n\mathbf{a}_n \to 0.$$

EXAMPLE 1 If $\mathfrak{S}$ is represented by interval $[1, 2]$ of $\mathbb{R}$, the following are symmetric deviation functions for $p > 0$: $|x - y|^p$, $\frac{|x-y|^p}{\min(x,y)}$. Examples of asymmetric deviations: $|x - y| \, e^{x-y}$, $|e^{y-x} - 1|$. Consider the latter function. Properties $\mathcal{D}1$ and $\mathcal{D}2$ being obvious, observe that $|e^{y-x} - 1|$ is continuous on a compactum, hence uniformly continuous: $\left|e^{b-a} - 1\right| - \left|e^{b'-a'} - 1\right| \to 0$ whenever $|a - a'| \to 0$ and $|b - b'| \to 0$. Property $\mathcal{D}3$ now follows from the observation that $|x - x'| \to 0$ if and only if $\left|e^{x-x'} - 1\right| \to 0$.

We now can define the notion of convergence induced by D on $\mathfrak{S}$.

DEFINITION 2 $\mathbf{a}_n \leftrightarrow \mathbf{b}_n$ iff $D\mathbf{a}_n\mathbf{b}_n \to 0$.

The notation is unambiguous because of the following theorem.

THEOREM 1 Convergence $\leftrightarrow$ is an equivalence relation.

Proof. Reflexivity is obvious. Symmetry is obtained by

$$(D\mathbf{a}_n\mathbf{b}_n \to 0) \wedge (D\mathbf{a}_n\mathbf{a}_n \to 0) \implies D\mathbf{a}_n\mathbf{a}_n - D\mathbf{b}_n\mathbf{a}_n \to 0.$$

Transitivity:

$$(D\mathbf{a}_n\mathbf{b}_n \to 0) \wedge (D\mathbf{b}_n\mathbf{c}_n \to 0) \implies (D\mathbf{a}_n\mathbf{b}_n \to 0) \wedge (D\mathbf{c}_n\mathbf{b}_n \to 0) \implies D\mathbf{a}_n\mathbf{c}_n - D\mathbf{b}_n\mathbf{b}_n \to 0.$$

∎

In particular, $\mathbf{a}_n \leftrightarrow \mathbf{a}$ means both $D\mathbf{a}\mathbf{a}_n \to 0$ and $D\mathbf{a}_n\mathbf{a} \to 0$ (because of which it is perfectly meaningful to write $\mathbf{a} \leftrightarrow \mathbf{a}_n$).

The convergence $\left(\mathbf{a}_n^1, ..., \mathbf{a}_n^k\right) \leftrightarrow \left(\mathbf{b}_n^1, ..., \mathbf{b}_n^k\right)$ can be defined by $\max_i D\mathbf{a}_n^i\mathbf{b}_n^i \to 0$.

2.2. Topology and Uniformity on $(\mathfrak{S}, D)$

A *topological basis* on $\mathfrak{S}$ is a family of subsets of $\mathfrak{S}$ satisfying the following property (Kelly, 1955, p. 47): if $\mathfrak{a}$ and $\mathfrak{b}$ are within the basis, then for any $\mathbf{x} \in \mathfrak{a} \cap \mathfrak{b}$ the basis contains a set $\mathfrak{c}$ which contains $\mathbf{x}$.

Given a topological basis on $\mathfrak{S}$, the *topology* on $\mathfrak{S}$ (a family of *open sets* "based" on this basis) is obtained by taking all possible unions of the subsets comprising the basis (including the empty set, which is the union of an empty class of such subsets).

THEOREM 2 Deviation D induces on $\mathfrak{S}$ a topology based on

$$\mathfrak{B}_D\left(\mathbf{x}, \varepsilon\right) = \{\mathbf{y} \in \mathfrak{S} : D\mathbf{x}\mathbf{y} < \varepsilon\}$$

taken for all $\mathbf{x} \in \mathfrak{S}$ and all real $\varepsilon > 0$.

Proof. We have to show that if $\mathbf{x} \in \mathfrak{B}_D\left(\mathbf{a},\alpha\right) \cap \mathfrak{B}_D\left(\mathbf{b},\beta\right)$, then one can find an $\varepsilon > 0$ such that $\mathfrak{B}_D\left(\mathbf{x},\varepsilon\right) \subset \mathfrak{B}_D\left(\mathbf{a},\alpha\right) \cap \mathfrak{B}_D\left(\mathbf{b},\beta\right)$. If, for every ε_n in a sequence $\varepsilon_n \to 0$ one could find an $\mathbf{x}_n \in \mathfrak{B}_D\left(\mathbf{x},\varepsilon_n\right)$ such that $\mathbf{x}_n \notin \mathfrak{B}_D\left(\mathbf{a},\alpha\right) \cap \mathfrak{B}_D\left(\mathbf{b},\beta\right)$, then we would have $\mathbf{x}_n \leftrightarrow \mathbf{x}$ but, for every n, $D\mathbf{a}\mathbf{x}_n \geq \alpha$ or $D\mathbf{b}\mathbf{x}_n \geq \beta$. This contradicts $\mathbf{x}_n \leftrightarrow \mathbf{x} \Longrightarrow \left(D\mathbf{a}\mathbf{x}_n \to D\mathbf{a}\mathbf{x} < \alpha\right) \wedge \left(D\mathbf{b}\mathbf{x}_n \to D\mathbf{b}\mathbf{x} < \beta\right)$. ∎

We call the topology just defined (based on $\mathfrak{B}_D$-balls) the *D*-topology.

THEOREM 3 The *D*-topology is also based on "reverse balls"

$$\mathfrak{B}_D^{-1}\left(\mathbf{x}, \varepsilon\right) = \{\mathbf{y} \in \mathfrak{S} : D\mathbf{y}\mathbf{x} < \varepsilon\}$$

taken for all $\mathbf{x} \in \mathfrak{S}$ and all real $\varepsilon > 0$.

Proof. That $\mathfrak{B}_D^{-1}\left(\mathbf{x}, \varepsilon\right)$ balls form a basis for a topology is shown by essentially repeating the proof of Theorem 2. That the bases $\mathfrak{B}_D^{-1}\left(\mathbf{x}, \varepsilon\right)$ and $\mathfrak{B}_D\left(\mathbf{x},\varepsilon\right)$ refine each other (i.e., every $\mathfrak{B}_D\left(\mathbf{x},\varepsilon\right)$ contains some $\mathfrak{B}_D^{-1}\left(\mathbf{x}, \delta\right)$, and vice versa)[9] follows from the symmetry of convergence (Theorem 1). ∎

COMMENT Although we do not need closed balls in this paper, it is useful to note that

$$\mathfrak{B}_D\left[\mathbf{x},\varepsilon\right] = \{\mathbf{y} \in \mathfrak{S} : D\mathbf{x}\mathbf{y} \leq \varepsilon\}$$

is a closed set, and $\overline{\mathfrak{B}}_D\left(\mathbf{x},\varepsilon\right) \subset \mathfrak{B}_D\left[\mathbf{x},\varepsilon\right]$ (where the overbar indicates topological closure). The two latter sets are not necessarily equal. Consider, e.g., $\mathfrak{S} = \mathbb{N}$, with $D\mathbf{x}\mathbf{y} = |x - y|$. Here, $\overline{\mathfrak{B}}_D\left(\mathbf{10},1\right) = \{\mathbf{10}\}$, while $\mathfrak{B}_D\left[\mathbf{10},1\right] = \{\mathbf{9}, \mathbf{10}, \mathbf{11}\}$.

Both previous theorems, as it turns out, can be strengthened: D induces on $\mathfrak{S}$ not only a topology but a more restrictive structure, called *uniformity*.

Recall (Kelly, 1955, p. 177), that a family of subsets of $\mathfrak{S} \times \mathfrak{S}$ forms a *basis for a uniformity* on $\mathfrak{S}$ if it satisfies the following four properties: if $\mathfrak{A}$ and $\mathfrak{B}$ are members of the basis, then

1. $\mathfrak{A}$ includes as its subset $\Delta = \{(\mathbf{x}, \mathbf{x}) : \mathbf{x} \in \mathfrak{S}\}$;

2. $\mathfrak{A}^{-1} = \{(\mathbf{y}, \mathbf{x}) : (\mathbf{y}, \mathbf{x}) \in \mathfrak{A}\}$ includes as its subset a member of the basis;

3. for some member $\mathfrak{C}$ of the basis, $\left\{(\mathbf{x}, \mathbf{z}) \in \mathfrak{S}^2 : \text{for some } \mathbf{y}, (\mathbf{x}, \mathbf{y}) \in \mathfrak{C} \wedge (\mathbf{y}, \mathbf{z}) \in \mathfrak{C}\right\} \subset \mathfrak{A}$;

4. $\mathfrak{A} \cap \mathfrak{B}$ includes as its subset a member of the basis.

Given a uniformity basis on $\mathfrak{S}$, the uniformity on $\mathfrak{S}$ ("based" on this basis) is obtained by taking each member of the basis and forming its unions with all subsets of $\mathfrak{S} \times \mathfrak{S}$. A member of a uniformity is called an *entourage*. A uniformity satisfies the separation axiom if the intersection of all its entourages is $\Delta = \{(\mathbf{x}, \mathbf{x}) : \mathbf{x} \in \mathfrak{S}\}$.

---

[9] If the proposition in the parentheses holds, then every $\mathfrak{B}_D\left(\mathbf{x},\varepsilon\right)$ contains a $\mathfrak{B}_D^{-1}\left(\mathbf{y}, \delta\right)$ for every $\mathbf{y} \in \mathfrak{B}_D\left(\mathbf{x},\varepsilon\right)$ (since $\mathfrak{B}_D\left(\mathbf{x},\varepsilon\right)$-balls form a basis), and vice versa. In general, a topological basis $\mathcal{A}$ is said to refine a topological basis $\mathcal{B}$ if for any $\mathbf{x} \in \mathfrak{a} \in \mathcal{A}$ there is a $\mathfrak{b} \in \mathcal{B}$ such that $\mathfrak{b} \subseteq \mathfrak{a}$. If $\mathcal{A}$ and $\mathcal{B}$ refine each other, they define one and the same topology.

THEOREM 4 *Deviation D induces on $\mathfrak{S}$ a uniformity based on entourages*

$$\mathfrak{U}_D(\varepsilon) = \left\{ (\mathbf{x}, \mathbf{y}) \in \mathfrak{S}^2 : D\mathbf{x}\mathbf{y} < \varepsilon \right\}$$

taken for all real $\varepsilon > 0$. The D-uniformity satisfies the separation axiom:

$$\cap_\varepsilon \mathfrak{U}_D(\varepsilon) = \left\{ (\mathbf{x}, \mathbf{y}) \in \mathfrak{S}^2 : \mathbf{x} = \mathbf{y} \right\}.$$

Proof. (1) $\Delta = \left\{ (\mathbf{x}, \mathbf{y}) \in \mathfrak{S}^2 : \mathbf{x} = \mathbf{y} \right\}$ is a subset of every $\mathfrak{U}_D(\varepsilon)$, obviously. (2) For any $\varepsilon > 0$ one can always find a $\delta > 0$ such that $\mathfrak{U}_D(\delta) \subset \mathfrak{U}_D^{-1}(\varepsilon) = \left\{ (\mathbf{x}, \mathbf{y}) \in \mathfrak{S}^2 : D\mathbf{y}\mathbf{x} < \varepsilon \right\}$, for otherwise we would be able to create sequences $\mathbf{x}_n \leftrightarrow \mathbf{y}_n$ with $D\mathbf{y}_n\mathbf{x}_n \geq \varepsilon$. (3) For any $\varepsilon > 0$ one can always find a $\delta > 0$ such that $(D\mathbf{x}\mathbf{y} < \delta) \wedge (D\mathbf{y}\mathbf{z} < \delta) \implies D\mathbf{x}\mathbf{z} < \varepsilon$, for otherwise we would have $\mathbf{x}_n \leftrightarrow \mathbf{y}_n$ and $\mathbf{y}_n \leftrightarrow \mathbf{z}_n$ but $D\mathbf{x}_n\mathbf{z}_n \geq \varepsilon$. (4) Finally, $\mathfrak{U}_D(\varepsilon) \cap \mathfrak{U}_D(\delta) = \mathfrak{U}_D(\min\{\delta, \varepsilon\})$. The last statement is obvious. ■

We call the uniformity just defined the D-uniformity. The D-topology is precisely the topology induced by the D-uniformity (Kelly, 1955, p. 178):

$$\mathfrak{B}_D(\mathbf{x}, \varepsilon) = \left\{ \mathbf{y} \in \mathfrak{S} : (\mathbf{x}, \mathbf{y}) \in \mathfrak{U}_D(\varepsilon) \right\}.$$

is the restriction of the basic entourage $\mathfrak{U}_D(\varepsilon)$ to the pairs $(\mathbf{x} = \mathbf{a}, \mathbf{y})$.

THEOREM 5 *The D-uniformity is also based on "reverse entourages"*

$$\mathfrak{U}_D^{-1}(\varepsilon) = \left\{ (\mathbf{x}, \mathbf{y}) \in \mathfrak{S}^2 : D\mathbf{y}\mathbf{x} < \varepsilon \right\}$$

taken for all real $\varepsilon > 0$.

Proof. That $\mathfrak{U}_D^{-1}(\varepsilon)$ entourages form a basis for a uniformity is shown by essentially repeating the proof of Theorem 4. That the bases $\mathfrak{U}_D^{-1}(\varepsilon)$ and $\mathfrak{U}_D(\varepsilon)$ refine each other thereby determining one and the same uniformity (i.e., each $\mathfrak{U}_D(\varepsilon)$ contains some $\mathfrak{U}_D^{-1}(\delta)$ and vice versa) follows from the symmetry of convergence (Theorem 1). ■

2.3. Chains

Finite chains in space $\mathfrak{S}$ are sequences of elements written as strings : $\mathbf{ab}$, $\mathbf{abc}$, $\mathbf{x}_1...\mathbf{x}_k$, etc. Note that the elements of a chain need not be pairwise distinct. A chain of *cardinality* k (a k-chain) is the chain with k elements (vertices), hence with $k-1$ links (edges). For completeness, we also admit an empty chain, of zero cardinality.

We use the notation

$$D\mathbf{x}_1...\mathbf{x}_k = \sum_{i=1}^{k-1} D\mathbf{x}_i\mathbf{x}_{i+1}. \tag{18}$$

and call it the D-length of the chain $\mathbf{x}_1...\mathbf{x}_k$.

If the elements of a chain are not of interest, it can be denoted by a boldface capital, such as $\mathbf{X}$, with appropriate ornaments. Thus, $\mathbf{X}$ and $\mathbf{Y}$ are two chains, $\mathbf{XY}$ is their concatenation, $\mathbf{aXb}$ is a chain connecting $\mathbf{a}$ to $\mathbf{b}$. The cardinality of chain $\mathbf{X}$ is denoted $|\mathbf{X}|$. Unless otherwise specified, within a sequence of chains, $\mathbf{X}_n$, the cardinality $|\mathbf{X}_n|$ generally varies: $\mathbf{X}_n = \mathbf{x}_1^n...\mathbf{x}_{k_n}^n$.

2.4.   Uniform Dissimilarity Function

DEFINITION 3 A uniform deviation function D on $\mathfrak{S}$ is a uniform dissimilarity (or, simply, dissimilarity) function on $\mathfrak{S}$ if it has the following property (see Fig. 9):

$\mathcal{D}4$. for any sequence of chains $\mathbf{a}_n\mathbf{X}_n\mathbf{b}_n$ with distinct elements and $|\mathbf{X}_n| \to \infty$,

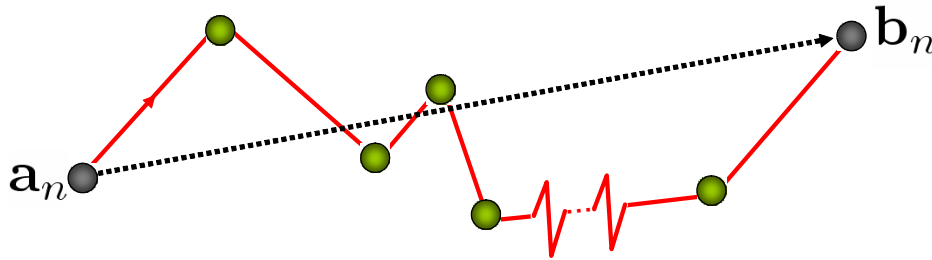$$D\mathbf{a}_n\mathbf{X}_n\mathbf{b}_n \to 0 \Longrightarrow D\mathbf{a}_n\mathbf{b}_n \to 0.$$

Figure 9. An Illustration for Property $\mathcal{D}4$. Consider an infinite sequence of chains $\mathbf{a}_1\mathbf{X}_1\mathbf{b}_1$, $\mathbf{a}_2\mathbf{X}_2\mathbf{b}_2$, ..., arbitrarily situated with respect to each other but such that $|\mathbf{X}_n|$ increases beyond bounds with $n \to \infty$, and $D\mathbf{a}_n\mathbf{X}_n\mathbf{b}_n$ gradually vanishes. Then $\mathbf{a}_n\mathbf{b}_n$ (the D-length of the dotted arrow) gradually vanishes too.

EXAMPLE 2 This example shows that $\mathcal{D}4$ does not follow from $\mathcal{D}1 - \mathcal{D}3$. Let $\mathfrak{S}$ be represented by $[1, 2]$ and, for $p > 1$, $D\mathbf{xy} = |x - y|^p$. This is a deviation function (Example 1). Consider the sequence of chains $\mathbf{aX}_n\mathbf{b}$ where *the* numerical values of $\mathbf{X}_n$ divide the interval $[a, b]$ into n equal parts. It is easy to see that $D\mathbf{aX}_n\mathbf{b} = n\left(\frac{b-a}{n}\right)^p \to 0$ while $D\mathbf{ab}$, obviously, remains equal to $(b - a)^p$.

The simple theorem below shows that the condition $|\mathbf{X}_n| \to \infty$ in the formulation of Property $\mathcal{D}4$ can be dropped.

THEOREM 6 Condition $\mathcal{D}4$ is equivalent to statements

$\mathcal{D}4a$. for any sequence of chains $\mathbf{a}_n\mathbf{X}_n\mathbf{b}_n$, $D\mathbf{a}_n\mathbf{X}_n\mathbf{b}_n \to 0 \Longrightarrow D\mathbf{a}_n\mathbf{b}_n \to 0$;

$\mathcal{D}4b$. for any $\varepsilon > 0$ one can find a $\delta > 0$ such that for any chain $\mathbf{aXb}$, if $D\mathbf{aXb} < \delta$, then $D\mathbf{ab} < \varepsilon$.

Proof. That $\mathcal{D}4b$ is merely a restatement of $\mathcal{D}4a$, and that $\mathcal{D}4a \implies \mathcal{D}4$ is obvious. It remains to show that $\mathcal{D}4 \implies \mathcal{D}4a$, even if $|\mathbf{X}_n| \not\to \infty$. Assume the contrary: D satisfies $\mathcal{D}4$ but there is a sequence $\mathbf{a}_n\mathbf{X}_n\mathbf{b}_n$ with $D\mathbf{a}_n\mathbf{X}_n\mathbf{b}_n \to 0$ and $D\mathbf{a}_n\mathbf{b}_n \not\to 0$. Then, for some subsequence $\mathbf{a}_{n_k}\mathbf{X}_{n_k}\mathbf{b}_{n_k}$ with distinct elements, we have $D\mathbf{a}_{n_k}\mathbf{X}_{n_k}\mathbf{b}_{n_k} \to 0$ and $D\mathbf{a}_{n_k}\mathbf{b}_{n_k} \to p > 0$ (finite or infinite). If $\limsup_{k\to\infty} |\mathbf{X}_{n_k}| = \infty$, then there must exist a subsequence $\mathbf{a}_{n_{k_l}}\mathbf{X}_{n_{k_l}}\mathbf{b}_{n_{k_l}}$ with $\left|\mathbf{X}_{n_{k_l}}\right| \to \infty$, $D\mathbf{a}_{n_{k_l}}\mathbf{X}_{n_{k_l}}\mathbf{b}_{n_{k_l}} \to 0$, but $D\mathbf{a}_{n_{k_l}}\mathbf{b}_{n_{k_l}} \to p > 0$ – which contradicts $\mathcal{D}4$. If $\limsup_{k\to\infty} |\mathbf{X}_{n_k}| < \infty$, then $|\mathbf{X}_{n_k}|$ are less than some natural N for all k, and one can redefine $\mathbf{X}_{n_k}$ to have $|\mathbf{X}_{n_k}| = N$, for all k (e.g., by replicating the last element of $\mathbf{X}_{n_k}$ a requisite number of times). The assumption $D\mathbf{a}_{n_k}\mathbf{X}_{n_k}\mathbf{b}_{n_k} \to 0$ then implies $\mathbf{a}_{n_k} \leftrightarrow \mathbf{x}_1^{n_k} \leftrightarrow ... \leftrightarrow \mathbf{x}_N^{n_k} \leftrightarrow \mathbf{b}_{n_k}$, and $\mathbf{a}_{n_k} \leftrightarrow \mathbf{b}_{n_k}$ is obtained by the transitivity of convergence (Theorem 1), in a contradiction to $D\mathbf{a}_{n_k}\mathbf{b}_{n_k} \to p > 0$. ∎

If D is a conventional metric or an oriented metric which is a deviation function (i.e., satisfies $D\mathbf{a}_n\mathbf{a}'_n \to 0 \implies D\mathbf{a}'_n\mathbf{a}_n \to 0$, as we know from Section 2.1), then D is a dissimilarity function as a trivial consequence of the triangle inequality:

$$D\mathbf{a}_n\mathbf{X}_n\mathbf{b}_n \geq D\mathbf{a}_n\mathbf{b}_n \geq 0$$

whence $D\mathbf{a}_n\mathbf{X}_n\mathbf{b}_n \to 0$ implies $D\mathbf{a}_n\mathbf{b}_n \to 0$.

EXAMPLE 3 In reference to the previous example, the same function $D\mathbf{xy} = |x - y|^p$ on $[1, 2]$ but with $0 < p \leq 1$ is a dissimilarity function because it is a metric. The function $D\mathbf{xy} = |e^{y-x} - 1|$ of Example 1 is not a metric, but it is an asymmetric dissimilarity function because $D\mathbf{a}_n\mathbf{X}_n\mathbf{b}_n$, as will be clear from the next section, never falls below $b_n - a_n$ if $a_n \leq b_n$ or below $1 - e^{b_n - a_n}$ if $a_n > b_n$. In either case $D\mathbf{a}_n\mathbf{X}_n\mathbf{b}_n \to 0$ implies $D\mathbf{a}_n\mathbf{b}_n = \left|e^{b_n - a_n} - 1\right| \to 0$.

The last example, with $D\mathbf{xy} = |e^{y-x} - 1|$, makes an implicit use of the oriented distance induced by D. This topic is taken on next.

2.5. Distance

The set of all possible chains in $\mathfrak{S}$ is denoted by $\mathcal{C}_{\mathfrak{S}}$, or simply $\mathcal{C}$. We define function $G\mathbf{ab}$ by

$$G\mathbf{ab} = \inf_{\mathbf{X}\in\mathcal{C}} D\mathbf{aXb}. \tag{19}$$

THEOREM 7 $G\mathbf{ab}$ is an oriented metric, and $G^*\mathbf{ab} = G\mathbf{ab} + G\mathbf{ba}$ is a metric (called overall, or symmetric).

Proof. Clearly, $G\mathbf{ab} \geq 0$. That $G\mathbf{aa} = 0$ is obvious. If $G\mathbf{ab} = 0$, then there is a sequence of chains $D\mathbf{aX}_n\mathbf{b} \to 0$, and this implies $\mathbf{a} = \mathbf{b}$ by Definition 3. Finally, the triangle inequality follows from the fact that the set of all chains $\mathbf{aXbYc}$ is a subset of the set of all chains $\mathbf{aZc}$. The statement about G^* is obvious. ∎

Due to this theorem we say that the oriented metric G and the overall (symmetric) metric G^* are induced by the dissimilarity D. Clearly, $G^*\mathbf{ab}$ can also be defined by (see Fig. 7)

$$G^*\mathbf{ab} = \inf_{(\mathbf{X},\mathbf{Y})\in\mathcal{C}^2} D\mathbf{aXbYa} = \inf_{(\mathbf{X},\mathbf{Y})\in\mathcal{C}^2} D\mathbf{bXaYb}. \tag{20}$$

EXAMPLE 4 Let $\mathfrak{S}$ be represented by $[1,2]$. Since $D\mathbf{xy} = |x-y|^p$ is a metric for $p \le 1$, we have $G\mathbf{xy} = D\mathbf{xy}$ and $G^*\mathbf{xy} = 2D\mathbf{xy}$. If $D\mathbf{xy} = \frac{|x-y|}{\min(x,y)}$, $G\mathbf{xy}$ can be shown to be $\left|\ln\frac{y}{x}\right|$. To outline the proof: we first show, by straightforward algebra, that for any $x < m < y$, $D\mathbf{xmy} < D\mathbf{xy}$ (i.e., $\frac{m-x}{x} + \frac{y-m}{m} < \frac{y-x}{x}$); it follows that the denser the sequence of intermediate points covering the interval between x and y the closer the D-length of the corresponding chain to the infimum;[10] the infimum therefore is the Riemann integral $\int_x^y \frac{(u+du)-u}{u} = \ln\frac{y}{x}$; symmetry considerations yield the formula for $x > y$. $G^*\mathbf{xy}$ here equals $2\left|\ln\frac{y}{x}\right|$. By analogous reasoning one can show that in the case $D\mathbf{xy} = |e^{y-x} - 1|$, for $x \le y$ the infimum is achieved at $G\mathbf{xy} = \int_x^y \left.\frac{d(e^{v-u}-1)}{dv}\right|_{v=u+} du = y - x$. For $x > y$, $D\mathbf{xy} = 1 - e^{y-x}$ is easily shown to be a metric, whence $G\mathbf{xy} = D\mathbf{xy} = 1 - e^{y-x}$. The overall (symmetric) distance in this case is $G^*\mathbf{xy} = |y-x| + 1 - e^{-|y-x|}$. In the "entomological" metaphor of Section 1.2 this example was discussed for $x = 1$, $y = 2$.

2.6. Topology and Uniformity on $(\mathfrak{S}, G)$

We proceed now to consider the topology and uniformity induced by the oriented metric G on $\mathfrak{S}$, the main result being that they coincide with those induced by the dissimilarity D which induces G.

THEOREM 8 $D\mathbf{a}_n\mathbf{b}_n \to 0 \iff G\mathbf{a}_n\mathbf{b}_n \to 0$

Proof. $D\mathbf{a}_n\mathbf{b}_n \to 0 \implies G\mathbf{a}_n\mathbf{b}_n \to 0$ follows from $0 \le G\mathbf{a}_n\mathbf{b}_n \le D\mathbf{a}_n\mathbf{b}_n$.

$G\mathbf{a}_n\mathbf{b}_n \to 0$ implies that for every n one can find a chain $\mathbf{X}_n$ such that

$$G\mathbf{a}_n\mathbf{b}_n \le D\mathbf{a}_n\mathbf{X}_n\mathbf{b}_n < G\mathbf{a}_n\mathbf{b}_n + \frac{1}{n}$$

whence $D\mathbf{a}_n\mathbf{X}_n\mathbf{b}_n \to 0$. Since D is a dissimilarity, this implies $D\mathbf{a}_n\mathbf{b}_n \to 0$. ∎

THEOREM 9 $\mathbf{a}_n \leftrightarrow \mathbf{b}_n \iff G\mathbf{a}_n\mathbf{b}_n \to 0 \iff G\mathbf{b}_n\mathbf{a}_n \to 0 \iff G^*\mathbf{a}_n\mathbf{b}_n = G^*\mathbf{b}_n\mathbf{a}_n \to 0$.

Proof. A simple consequence of Theorems 8 and 1. ∎

THEOREM 10 G induces on $\mathfrak{S}$ a topology based on any of the three following sets: the set of

$$\mathfrak{B}_G(\mathbf{x},\varepsilon) = \{\mathbf{y} \in \mathfrak{S} : G\mathbf{xy} < \varepsilon\},$$

---

[10]This is an example of an important special case in the DC theory, a space "with intermediate points." Such spaces are arc-connected, and with the length of continuous paths appropriately defined, the paths replace finite chains in the computation of distances. The specialization of the DC theory to arc-connected spaces is considered in Dzhafarov (2007).

the set of

$$\mathfrak{B}_G^{-1}(\mathbf{x}, \varepsilon) = \{\mathbf{y} \in \mathfrak{S} : G\mathbf{y}\mathbf{x} < \varepsilon\},$$

and the set of

$$\mathfrak{B}_{G^*}(\mathbf{x}, \varepsilon) = \{\mathbf{y} \in \mathfrak{S} : G^*\mathbf{x}\mathbf{y} < \varepsilon\},$$

taken for all $\mathbf{x} \in \mathfrak{S}$ and positive ε. This topology coincides with the D-topology.

Proof. A simple consequence of Theorems 8 and 9. ∎

THEOREM 11 G induces on $\mathfrak{S}$ a uniformity based on on any of the three following sets: the set of

$$\mathfrak{U}_G(\varepsilon) = \{(\mathbf{x}, \mathbf{y}) \in \mathfrak{S}^2 : G\mathbf{x}\mathbf{y} < \varepsilon\},$$

the set of

$$\mathfrak{U}_G^{-1}(\varepsilon) = \{(\mathbf{x}, \mathbf{y}) \in \mathfrak{S}^2 : G\mathbf{y}\mathbf{x} < \varepsilon\},$$

and set of

$$\mathfrak{U}_{G^*}(\varepsilon) = \{(\mathbf{x}, \mathbf{y}) \in \mathfrak{S}^2 : G^*\mathbf{x}\mathbf{y} < \varepsilon\}$$

taken for all positive ε. This uniformity coincides with the D-uniformity.

Proof. Again, follows from Theorems 8 and 9 on observing that $\mathfrak{U}_G(\varepsilon)$, $\mathfrak{U}_G^{-1}(\varepsilon)$, and $\mathfrak{U}_{G^*}(\varepsilon)$ are standard metric entourages. ∎

The space $(\mathfrak{S}, D)$ being uniform and metrizable, we get its standard topological characterization (see, e.g., Hocking &Young, 1961, p. 42): it is a *completely normal space*, meaning that its singletons are closed and any its two separated subsets $\mathfrak{A}$ and $\mathfrak{B}$ (i.e., such that $\overline{\mathfrak{A}} \cap \mathfrak{B} = \mathfrak{A} \cap \overline{\mathfrak{B}} = \varnothing$) are contained in two disjoint open subsets. This characterization is much stronger than the Urysohn property established in Dzhafarov & Colonius (2005a): a completely normal space is Urysohn (hence also Hausdorff).

In conclusion we establish two basic properties of the oriented metric G: its uniform continuity and its intrinsicality with respect to our standard procedure (all possible chains, infimum of D-lengths).

THEOREM 12 $G\mathbf{x}\mathbf{y}$ is uniformly continuous in $(\mathbf{x}, \mathbf{y})$, i.e., if $\mathbf{a}'_n \leftrightarrow \mathbf{a}_n$ and $\mathbf{b}'_n \leftrightarrow \mathbf{b}_n$, then $G\mathbf{a}'_n\mathbf{b}'_n - G\mathbf{a}_n\mathbf{b}_n \to 0$.

Proof. Follows from $G\mathbf{a}'_n\mathbf{b}'_n \leq G\mathbf{a}'_n\mathbf{a}_n + G\mathbf{a}_n\mathbf{b}_n + G\mathbf{b}_n\mathbf{b}'_n$, and $G\mathbf{a}_n\mathbf{b}_n \leq G\mathbf{a}_n\mathbf{a}'_n + G\mathbf{a}'_n\mathbf{b}'_n + G\mathbf{b}'_n\mathbf{b}_n$ (by the triangle inequality), using the symmetry of convergence to zero (Theorem 9). ∎

The following is an important fact which can be interpreted as that of internal consistency of the metric G induced by means of our standard construction: once $G\mathbf{ab}$ is computed as the infimum of the D-length across all chains from $\mathbf{a}$ to $\mathbf{b}$, the infimum of the G-length across all chains from $\mathbf{a}$ to $\mathbf{b}$ equals $G\mathbf{ab}$. By analogy

with the traditional terminology, this means that G is an *intrinsic metric*.[11] The notation for cumulated G-length is analogous to that for D-length:

$$G\mathbf{x}_1...\mathbf{x}_k = \sum_{i=1}^{k-1} G\mathbf{x}_i\mathbf{x}_{i+1}. \tag{21}$$

THEOREM 13 $D\mathbf{aX}_n\mathbf{b} \to G\mathbf{ab} \implies G\mathbf{aX}_n\mathbf{b} \to G\mathbf{ab}$.

Proof. $G\mathbf{aX}_n\mathbf{b} \leq D\mathbf{aX}_n\mathbf{b}$, so $\limsup_{n\to\infty} G\mathbf{aX}_n\mathbf{b} \leq G\mathbf{ab}$. On the other hand, $G\mathbf{aX}_n\mathbf{b} \geq G\mathbf{ab}$ (triangle inequality), whence $\liminf_{n\to\infty} G\mathbf{aX}_n\mathbf{b} \geq G\mathbf{ab}$. ■

The reverse is obviously not true: thus, choosing any $\mathbf{a}, \mathbf{b}$ for which $G\mathbf{ab} < D\mathbf{ab}$ and any $\mathbf{x}_n \leftrightarrow \mathbf{a}$, we have (by the uniform continuity of both G and D) $G\mathbf{ax}_n\mathbf{b} \to G\mathbf{ab}$, but $D\mathbf{ax}_n\mathbf{b} \to D\mathbf{ab} > G\mathbf{ab}$.

2.7.   A Comment on Weak Dissimilarity Functions

The main function of Property $\mathcal{D}4$ ($\mathcal{D}4a$, $\mathcal{D}4b$) is to ensure that $G\mathbf{ab}$ for distinct $\mathbf{a}, \mathbf{b}$ never vanishes (Theorem 7). This goal alone, however, can be achieved more directly, by means of a weaker constraint.

DEFINITION 4 A uniform deviation function D on $\mathfrak{S}$ is a weak dissimilarity function on $\mathfrak{S}$ if it has the following property:

$\mathcal{D}4'$. for any sequence of chains $\mathbf{aX}_n\mathbf{b}$ with distinct elements and $|\mathbf{X}_n| \to \infty$,

$$D\mathbf{aX}_n\mathbf{b} \to 0 \implies D\mathbf{ab} = 0.$$

Here too we have two equivalent reformulations of the defining property:

$\mathcal{D}4'a$. for any sequence of chains $\mathbf{aX}_n\mathbf{b}$, $D\mathbf{aX}_n\mathbf{b} \to 0 \implies D\mathbf{ab} = 0$ (stipulation $|\mathbf{X}_n| \to \infty$ not necessary);

$\mathcal{D}4'b$. for any $\mathbf{a}, \mathbf{b}$, and $\varepsilon > 0$ one can find a $\delta > 0$ such that for any chain $\mathbf{X}$, if $D\mathbf{aXb} < \delta$, then $D\mathbf{ab} < \varepsilon$ (compare to $\mathcal{D}4b$, uniformity lost).

A version of Definition 4 was adopted in our preliminary exposition of the DC theory (Dzhafarov & Colonius, 2005c).[12] Some of the results given in this paper, as well as many results pertaining to special cases of the DC theory (to be presented separately), only require the weak form of a dissimilarity function.

---

[11] Traditionally the notion requires the existence of paths (continuous images of segments of reals) and their lengths (see, e.g., Dzhafarov & Colonius, 2001, 2005a), so our present usage is not standard. There are, however, special cases of DC where the infima of the D-lengths taken across a certain class of paths coincide with those taken across all finite chains. In such cases (considered in Dzhafarov, 2007) the intrinsicality of G acquires its traditional meaning.

[12] Dissimilarity functions with bounded codomains were defined in that work by Properties $\mathcal{D}1$, $\mathcal{D}2$, $\mathcal{D}3$, and (an equivalent of) $\mathcal{D}4'$. The treatment of unbounded codomains there was erroneous.

We do lose, however, the equivalence of the topologies and uniformities induced by D and by G (the last statements of Theorems 10 and 11), due to the fact that with a weak dissimilarity we only have

$$D\mathbf{a}_n\mathbf{b}_n \to 0 \Longrightarrow G\mathbf{a}_n\mathbf{b}_n \to 0$$

rather than the equivalence stated in Theorem 8. There are special cases (e.g., spaces compact in D-topology, wherein any infinite sequence contains a D-converging subsequence) when a weak dissimilarity is always a uniform dissimilarity too (i.e., $\mathcal{D}4$ and $\mathcal{D}4'$ are equivalent). As the following example shows, however, $\mathcal{D}4'$ generally does not imply $\mathcal{D}4$.
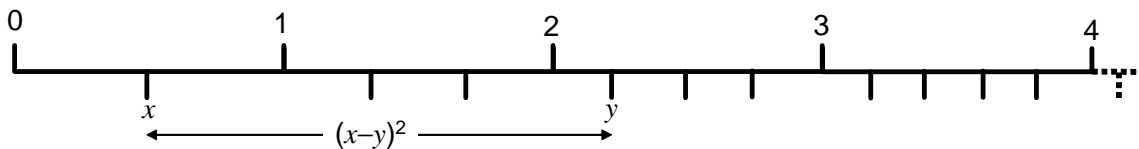

Figure 10. A space representation in which $\mathcal{D}4'$ holds but $\mathcal{D}4$ does not. The representation consists of isolated points shown by ticks on the number line. The dissimilarity function is $D\mathbf{xy} = (x - y)^2$.

EXAMPLE 5 (See Fig. 10.) Consider a discrete space numerically represented by all rational numbers of the form $n - 1 + \frac{k_n}{n+1}$, $n = 1, 2, \ldots$ and $k_n = 0, 1, \ldots, n$. In other words, the numbers include all nonnegative integers and the intermediate points dividing every interval $[n-1, n]$ into $n + 1$ equal parts. Let the dissimilarity function be $D\mathbf{xy} = (x - y)^2$. It is easy to see that the geodesic (shortest) chain exists between any two distinct points and consists of all points in between. The D-length of such a geodesic is clearly nonzero, and we conclude that $\mathcal{D}4'$ is satisfied. Consider, however, the sequence $\mathbf{a}_n\mathbf{b}_n$ with numerical values $(0, 1), (1, 2), \ldots, (n - 1, n), \ldots$. Denoting by $\mathbf{S}_n$ the geodesic chain between $\mathbf{a}_n$ and $\mathbf{b}_n$, we have $D\mathbf{a}_n\mathbf{S}_n\mathbf{b}_n = \frac{1}{n+1} \to 0$. At the same time, $D\mathbf{a}_n\mathbf{b}_n = 1$ for all n. We conclude that $\mathcal{D}4$ is not satisfied.

2.8. Universal Fechnerian Scaling

We are ready now to introduce UFS, the application of the DC theory to discrimination probabilities. The material presented in Section 1.3 is an integral part of UFS. We briefly recapitulate the main points.

1. We start with stimuli in two observation areas, $(\mathfrak{S}_1^*, \mathfrak{S}_2^*, \psi^*)$.

2. We "lump together" (label identically) psychologically equal stimuli, obtaining $\left(\mathfrak{S}_1, \mathfrak{S}_2, \tilde{\psi}\right)$.

3. We postulate the law of Regular Minimality.

4. We subject $\mathfrak{S}_1, \mathfrak{S}_2$ to a canonical transformation to get a canonical space $(\mathfrak{S}, \psi)$.

5. We define (canonical) psychometric increments of the first and second kind, $\Psi^{(1)}\mathbf{ab} = \psi\mathbf{ab} - \psi\mathbf{aa}$, and $\Psi^{(2)}\mathbf{ab} = \psi\mathbf{ba} - \psi\mathbf{aa}$.

We now add to the law of Regular Minimality the only other postulate of UFS:

$$\Psi^{(1)} \text{ and } \Psi^{(2)} \text{ are uniform dissimilarity functions.}$$

The application of the DC theory now yields functions G_1, G_2 defined by

$$G_1\mathbf{ab} = \inf_{\mathbf{X}\in\mathcal{C}} \Psi^{(1)}\mathbf{aXb}$$
$$G_2\mathbf{ab} = \inf_{\mathbf{X}\in\mathcal{C}} \Psi^{(2)}\mathbf{aXb} \tag{22}$$

and called the *oriented Fechnerian metrics* on $\mathfrak{S}$ (recall that $\mathcal{C}$ denotes the set of all finite chains in $\mathfrak{S}$). The notation $\Psi^{(1)}\mathbf{aXb}$, $\Psi^{(2)}\mathbf{aXb}$ has the same meaning as $D\mathbf{aXb}$, since $\Psi^{(1)}$, $\Psi^{(2)}$ are merely special versions of D: with $\mathbf{X} = \mathbf{x}_1...\mathbf{x}_k$,

$$\Psi^{(1)}\mathbf{aXb} = \Psi^{(1)}\mathbf{ax}_1 + \Psi^{(1)}\mathbf{x}_1\mathbf{x}_2 + ... + \Psi^{(1)}\mathbf{x}_k\mathbf{b}$$
$$\Psi^{(2)}\mathbf{aXb} = \Psi^{(2)}\mathbf{ax}_1 + \Psi^{(2)}\mathbf{x}_1\mathbf{x}_2 + ... + \Psi^{(2)}\mathbf{x}_k\mathbf{b}. \tag{23}$$

The function

$$G^*\mathbf{ab} = G_1\mathbf{ab} + G_1\mathbf{ba} = G_2\mathbf{ab} + G_2\mathbf{ba} \tag{24}$$

is called the *overall (symmetric) Fechnerian metric* on $\mathfrak{S}$. It can also be presented as

$$G^*\mathbf{ab} = \inf_{(\mathbf{X},\mathbf{Y})\in\mathcal{C}^2} \Psi^{(1)}\mathbf{aXbYa} = \inf_{(\mathbf{X},\mathbf{Y})\in\mathcal{C}^2} \Psi^{(2)}\mathbf{aXbYa}. \tag{25}$$

A formal proof for the equality of the two infima is straightforward.

THEOREM 14 For any stimuli $\mathbf{a}, \mathbf{b}$,

$$\inf_{(\mathbf{X},\mathbf{Y})\in\mathcal{C}^2} \Psi^{(1)}\mathbf{aXbYa} = \inf_{(\mathbf{X},\mathbf{Y})\in\mathcal{C}^2} \Psi^{(2)}\mathbf{aXbYa}.$$

Proof. For any two chains $\mathbf{x}_1\mathbf{x}_2...\mathbf{x}_{k-1}\mathbf{x}_k$ and $\mathbf{y}_1\mathbf{y}_2...\mathbf{y}_{l-1}\mathbf{y}_l$, consider the closed chains

$$\mathbf{L}_1 \quad = \quad \mathbf{ax}_1\mathbf{x}_2...\mathbf{x}_k\mathbf{by}_l\mathbf{y}_{l-1}...\mathbf{y}_1\mathbf{a}$$
$$\mathbf{L}_2 \quad = \quad \mathbf{ay}_1\mathbf{y}_2...\mathbf{y}_l\mathbf{bx}_k\mathbf{x}_{k-1}...\mathbf{x}_1\mathbf{a},$$

i.e., $\mathbf{L}_1$ and $\mathbf{L}_2$ include the same elements but traversed in opposite directions. Denoting $\mathbf{x}_0 = \mathbf{y}_0 = \mathbf{a}$ and $\mathbf{x}_{k+1} = \mathbf{y}_{l+1} = \mathbf{b}$,

$$\Psi^{(1)}\mathbf{L}_1 \quad = \quad \sum_{i=0}^{k} \Psi^{(1)}\mathbf{x}_i\mathbf{x}_{i+1} + \sum_{i=0}^{l} \Psi^{(1)}\mathbf{y}_{i+1}\mathbf{y}_i = \sum_{i=0}^{k}(\psi\mathbf{x}_i\mathbf{x}_{i+1} - \psi\mathbf{x}_i\mathbf{x}_i) + \sum_{i=0}^{l}(\psi\mathbf{y}_{i+1}\mathbf{y}_i - \psi\mathbf{y}_{i+1}\mathbf{y}_{i+1})$$

$$\Psi^{(2)}\mathbf{L}_2 \quad = \quad \sum_{i=0}^{k} \Psi^{(2)}\mathbf{x}_{i+1}\mathbf{x}_i + \sum_{i=0}^{l} \Psi^{(2)}\mathbf{y}_i\mathbf{y}_{i+1} = \sum_{i=0}^{k}(\psi\mathbf{x}_i\mathbf{x}_{i+1} - \psi\mathbf{x}_{i+1}\mathbf{x}_{i+1}) + \sum_{i=0}^{l}(\psi\mathbf{y}_{i+1}\mathbf{y}_i - \psi\mathbf{y}_i\mathbf{y}_i)$$

and

$$\Psi^{(1)}\mathbf{L}_1 - \Psi^{(2)}\mathbf{L}_2 \quad = \quad \sum_{i=0}^{k}(\psi\mathbf{x}_{i+1}\mathbf{x}_{i+1} - \psi\mathbf{x}_i\mathbf{x}_i) + \sum_{i=0}^{l}(\psi\mathbf{y}_i\mathbf{y}_i - \psi\mathbf{y}_{i+1}\mathbf{y}_{i+1})$$
$$= \quad (\psi\mathbf{x}_{k+1}\mathbf{x}_{k+1} - \psi\mathbf{x}_0\mathbf{x}_0) + (\psi\mathbf{y}_0\mathbf{y}_0 - \psi\mathbf{y}_{l+1}\mathbf{y}_{l+1})$$
$$= \quad (\mathbf{b} - \mathbf{a}) + (\mathbf{a} - \mathbf{b}) = 0$$

Now, the set of all possible pairs of chains $(\mathbf{x}_1\mathbf{x}_2...\mathbf{x}_k, \mathbf{y}_l\mathbf{y}_{l-1}...\mathbf{y}_1)$, for all k and l, is the same as the set of all possible pairs of chains $(\mathbf{y}_1\mathbf{y}_2...\mathbf{y}_l, \mathbf{x}_k\mathbf{x}_{k-1}...\mathbf{x}_1)$, for all k and l: in both cases the set is $\mathcal{C}^2$. Whence the statement of the theorem follows immediately. $\blacksquare$

We now proceed to establish the basic topological properties of the discrimination probability function.

THEOREM 15 $\Psi^{(1)}\mathbf{a}_n\mathbf{b}_n \to 0$ iff $\Psi^{(2)}\mathbf{a}_n\mathbf{b}_n \to 0$.

Proof. Follows from

$$\Psi^{(1)}\mathbf{a}_n\mathbf{b}_n + \Psi^{(1)}\mathbf{b}_n\mathbf{a}_n = \Psi^{(2)}\mathbf{a}_n\mathbf{b}_n + \Psi^{(2)}\mathbf{b}_n\mathbf{a}_n$$

and the fact that, for $\iota = 1, 2$,

$$\Psi^{(\iota)}\mathbf{a}_n\mathbf{b}_n \to 0 \iff \Psi^{(\iota)}\mathbf{b}_n\mathbf{a}_n \to 0.$$

$\blacksquare$

Because of this theorem we can write $\mathbf{a}_n \leftrightarrow \mathbf{b}_n$ unambiguously, meaning any of (and therefore all of) the relations

$$\Psi^{(1)}\mathbf{a}_n\mathbf{b}_n \to 0 \iff \psi\mathbf{a}_n\mathbf{b}_n - \psi\mathbf{a}_n\mathbf{a}_n \to 0$$
$$\Psi^{(1)}\mathbf{b}_n\mathbf{a}_n \to 0 \iff \psi\mathbf{b}_n\mathbf{a}_n - \psi\mathbf{b}_n\mathbf{b}_n \to 0$$
$$\Psi^{(2)}\mathbf{a}_n\mathbf{b}_n \to 0 \iff \psi\mathbf{b}_n\mathbf{a}_n - \psi\mathbf{a}_n\mathbf{a}_n \to 0$$
$$\Psi^{(2)}\mathbf{b}_n\mathbf{a}_n \to 0 \iff \psi\mathbf{a}_n\mathbf{b}_n - \psi\mathbf{b}_n\mathbf{b}_n \to 0$$

THEOREM 16 Function $\psi\mathbf{ab}$ is uniformly continuous: if $\mathbf{a}'_n \leftrightarrow \mathbf{a}_n$ and $\mathbf{b}'_n \leftrightarrow \mathbf{b}_n$, then $\psi\mathbf{a}'_n\mathbf{b}'_n - \psi\mathbf{a}_n\mathbf{b}_n \to 0$.

Proof. Since $\mathbf{a}'_n \leftrightarrow \mathbf{a}_n$ means both

$$\Psi^{(1)}\mathbf{a}_n\mathbf{a}'_n = \psi\mathbf{a}_n\mathbf{a}'_n - \psi\mathbf{a}_n\mathbf{a}_n \to 0$$

and

$$\Psi^{(2)}\mathbf{a}'_n\mathbf{a}_n = \psi\mathbf{a}_n\mathbf{a}'_n - \psi\mathbf{a}'_n\mathbf{a}'_n \to 0,$$

it implies

$$\psi\mathbf{a}'_n\mathbf{a}'_n - \psi\mathbf{a}_n\mathbf{a}_n \to 0.$$

Consequently $\mathbf{a}'_n \leftrightarrow \mathbf{a}_n$ and $\mathbf{b}'_n \leftrightarrow \mathbf{b}_n$ imply

$$\begin{aligned}
\psi\mathbf{a}'_n\mathbf{b}'_n - \psi\mathbf{a}_n\mathbf{b}_n &= \left[\Psi^{(1)}\mathbf{a}'_n\mathbf{b}'_n + \psi\mathbf{a}'_n\mathbf{a}'_n\right] - \left[\Psi^{(1)}\mathbf{a}_n\mathbf{b}_n + \psi\mathbf{a}_n\mathbf{a}_n\right] \\
&= \left[\Psi^{(1)}\mathbf{a}'_n\mathbf{b}'_n - \Psi^{(1)}\mathbf{a}_n\mathbf{b}_n\right] + [\psi\mathbf{a}'_n\mathbf{a}'_n - \psi\mathbf{a}_n\mathbf{a}_n] \to 0.
\end{aligned}$$

$\blacksquare$

3.   Conclusion

Universal Fechnerian Scaling is a theory dealing with the computation of subjective distances from pairwise discrimination probabilities. The theory is applicable to all possible stimulus spaces subject to the assumptions that

(A) discrimination probabilities satisfy the law of Regular Minimality, and

(B) the two canonical psychometric increments of the first and second kind, $\Psi^{(1)}$ and $\Psi^{(2)}$, are dissimilarity functions.

A dissimilarity function $D\mathbf{ab}$ (where D can stand for either $\Psi^{(1)}$ or $\Psi^{(2)}$) for pairs of stimuli in a canonical representation is defined by the following properties:

$\mathcal{D}1$. $\mathbf{a} \neq \mathbf{b} \Longrightarrow D\mathbf{ab} > 0$;

$\mathcal{D}2$. $D\mathbf{aa} = 0$;

$\mathcal{D}3$. If $D\mathbf{a}_n\mathbf{a}_n' \to 0$ and $D\mathbf{b}_n\mathbf{b}_n' \to 0$, then $D\mathbf{a}_n'\mathbf{b}_n' - D\mathbf{a}_n\mathbf{b}_n \to 0$; and

$\mathcal{D}4$. for any sequence $\mathbf{a}_n\mathbf{X}_n\mathbf{b}_n$, where $\mathbf{X}_n$ is a chain of stimuli, $D\mathbf{a}_n\mathbf{X}_n\mathbf{b}_n \to 0 \Longrightarrow D\mathbf{a}_n\mathbf{b}_n \to 0$.

The overall (symmetric) Fechnerian distance $G^*\mathbf{ab}$ between $\mathbf{a}$ and $\mathbf{b}$ is defined as the infimum of $D\mathbf{aXb}+D\mathbf{bYa}$ across all possible chains $\mathbf{X}$ and $\mathbf{Y}$ inserted between $\mathbf{a}$ and $\mathbf{b}$. This computation does not depend on whether one uses $\Psi^{(1)}$ or $\Psi^{(2)}$ in place of D.

The canonical psychometric increments $\Psi^{(1)}$ or $\Psi^{(2)}$ impose on stimulus space one and the same topology and uniformity structure, which also coincide with the topology and uniformity induced by the Fechnerian metric G^*. The discrimination probability function is uniformly continuous with respect to the uniformity just mentioned. Stimulus space is topologically characterized as a completely normal space.

Some of the important special cases of the DC theory and the corresponding cases of UFS will be discussed in follow-up papers. This prominently includes the specialization of DC/UFS to the previously published Multidimensional Fechnerian Scaling, and more general cases where the logic of DC leads us from finite connecting chains to continuous connecting paths.

References

Dzhafarov E.N. (2001a). Fechnerian Psychophysics. In N.J. Smelser, P.B. Baltes (Eds.) *International Encyclopedia of the Social and Behavioral Sciences*, v. 8 (pp. 5437-5440). New York: Pergamon Press.

Dzhafarov, E.N. (2001b). Fechnerian scaling and Thurstonian modeling. In E. Sommerfeld, R. Kompass, & T. Lachmann (Eds.), *Fechner Day 2001* (pp. 42-47). Lengerich: Pabst Science Publishers.

Dzhafarov, E.N. (2002a). Multidimensional Fechnerian scaling: Regular variation version. *Journal of Mathematical Psychology*, 46, 226-244.

Dzhafarov, E.N. (2002b). Multidimensional Fechnerian scaling: Probability-distance hypothesis. *Journal of Mathematical Psychology*, 46, 352-374.

Dzhafarov, E.N. (2002c). Multidimensional Fechnerian scaling: Perceptual separability. *Journal of Mathematical Psychology*, 46, 564-582.

Dzhafarov, E.N. (2002d). Multidimensional Fechnerian scaling: Pairwise comparisons, regular minimality, and nonconstant self-similarity. *Journal of Mathematical Psychology*, 46, 583-608.

Dzhafarov, E.N. (2003a). Thurstonian-type representations for "same-different" discriminations: Deterministic decisions and independent images. *Journal of Mathematical Psychology*, 47, 208-228.

Dzhafarov, E.N. (2003b). Thurstonian-type representations for "same-different" discriminations: Probabilistic decisions and interdependent images. *Journal of Mathematical Psychology*, 47, 229-243. [see Dzhafarov, E.N. (2006). Corrigendum to "Thurstonian-type representations for 'same–different' discriminations: Probabilistic decisions and interdependent images." *Journal of Mathematical Psychology*, 50, 511.]

Dzhafarov, E.N. (2004). Perceptual separability of stimulus dimensions: A Fechnerian approach. In C. Kaernbach, E. Schröger, H. Müller (Eds.), *Psychophysics beyond Sensation: Laws and Invariants of Human Cognition* (pp. 9-26). Mahwah, NJ: Erlbaum.

Dzhafarov, E.N. (2006). On the law of Regular Minimality: Reply to Ennis. *Journal of Mathematical Psychology*, 50, 74-93.

Dzhafarov, E.N. (2007). Dissimilarity Cumulation theory in arc-connected spaces. *Journal of Mathematical Psychology*, in press.

Dzhafarov, E.N., & Colonius, H. (1999a). Fechnerian metrics in unidimensional and multidimensional stimulus spaces. *Psychonomic Bulletin and Review*, 6, 239-268.

Dzhafarov, E.N., & Colonius, H. (1999b). Fechnerian metrics. In P.R. Kileen & W.R. Uttal (Eds.), *Looking Back: The End of the 20th Century Psychophysics* (pp. 111-116). Tempe, AZ: Arizona University Press.

Dzhafarov, E.N., & Colonius, H. (2001). Multidimensional Fechnerian scaling: Basics. *Journal of Mathematical Psychology*, 45, 670-719.

Dzhafarov, E.N., & Colonius, H. (2005a). Psychophysics without physics: A purely psychological theory of Fechnerian Scaling in continuous stimulus spaces. *Journal of Mathematical Psychology*, 49, 1-50.

Dzhafarov, E.N., & Colonius, H. (2005b). Psychophysics without physics: Extension of Fechnerian Scaling from continuous to discrete and discrete-continuous stimulus spaces. *Journal of Mathematical Psychology,* 49, 125-141.

Dzhafarov, E.N., & Colonius, H. (2005c). From Fechner to Fechner, and beyond. In J.S. Monahan, S.M. Sheffert, & J.T. Townsend (Eds) *Fechner Day 2005* (pp. 83-88). Mt. Pleasant, MI: Michgan University Press.

Dzhafarov, E.N., & Colonius, H. (2006a). Regular Minimality: A fundamental law of discrimination. In H. Colonius & E.N. Dzhafarov (Eds.), *Measurement and Representation of Sensations* (pp. 1-46). Mahwah, NJ: Erlbaum.

Dzhafarov, E.N., & Colonius, H. (2006b). Generalized Fechnerian Scaling. In H. Colonius & E.N. Dzhafarov (Eds.), *Measurement and Representation of Sensations* (pp. 47-88). Mahwah, NJ: Erlbaum.

Dzhafarov, E.N., & Colonius, H. (2006c). Reconstructing distances among objects from their discriminability. *Psychometrika,* 71, 365 - 386.

Ennis, D. M. (2006). Sources and influence of perceptual variance: Comment on Dzhafarov's Regular Minimality Principle. *Journal of Mathematical Psychology,* 50, 66–73.

Hocking, J.H. &Young G.S. (1961). *Topology.* Reading, MA: Addison-Wesley.

Iverson, G.J. (2006). Analytical methods in the theory of psychophysical discrimination I: Inequalities, convexity and integration of just noticeable differences. *Journal of Mathematical Psychology,* 50, 271-282.

Kelly, J.L. (1955). *General Topology.* Toronto: Van Nostrand.

Townsend, J.T., Aisbett, J., Busemeyer, J., & Assadi, A. (2006). Methodology for dimensional independence on simple cognitive manifolds. In H. Colonius & E.N. Dzhafarov (Eds.), *Measurement and Representation of Sensations* (pp. 203-241). Mahwah, NJ: Erlbaum.

Zhang, J. (2004). Dual scaling between comparison and reference stimuli in multidimensional psychological space. *Journal of Mathematical Psychology,* 48: 409-424.

Zhang, J. (2006). Referential duality and representational duality in the scaling of multi-dimensional and infinite-dimensional stimulus space. In H. Colonius & E.N. Dzhafarov (Eds.), *Measurement and Representation of Sensations* (pp. 131-157). Mahwah, NJ: Erlbaum.