ROCZNIK NAUKOWO-DYDAKTYCZNY AKADEMII PEDAGOGICZNEJ W KRAKOWIE

Zeszyt 204 Prace Matematyczne XVII 2000

Zoltán Boros

Characterization of transformed linear functions via shift invariances

Dedicated to Professor Zenon Moszner
on his 70th birthday

Abstract. Transforms of linear functions in two variables are characterized via translation like composite functional equations on restricted domain, under regularity assumptions of $\mathcal{C}^1$ type.

1. Introduction

In psychophysical theories of binocular space perception various formulae are established for perceived egocentric distance and perceived egocentric direction in bipolar coordinates (i.e. in terms of the monocular directions relative to the rotation centers of the eyes) in the horizontal plane at eye-level. These formulae are usually based on sets of experimental data but they are often presented bare of any theoretical justification. A geometrically motivated theory of binocular space perception has been suggested by Luneburg [3]. In his model the perceived egocentric distance $\rho(\alpha,\beta) > 0$ and the perceived direction $\vartheta(\alpha,\beta) \in \left]-\frac{\pi}{2}, \frac{\pi}{2}\right[$, defined on

$$S = \left\{(\alpha,\beta) \in \mathbb{R}^2 \mid -\frac{\pi}{2} < \beta < \alpha < \frac{\pi}{2}\right\},$$

are expressed by

$$\rho(\alpha,\beta) = \alpha - \beta \quad \text{and} \quad \vartheta(\alpha,\beta) = \frac{\alpha+\beta}{2}.$$

Mathematics Subject Classification (2000): Primary 39B22, Secondary 39B12, 26B40.

The author's research was supported by the grants OTKA T-030082, FKFP 0310/1997 and TDFK (Debrecen).

From these formulae, using basic arguments of elementary and analytic geometry, one can derive that, in Luneburg's model, the level sets of points with equal perceived egocentric distances are circles going through the centers of the eyes, while those of points with equal perceived egocentric directions are hyperbolae, which are asymptotically close to half lines with the appropriate physical direction.

In a recent attempt to give a qualitative characterization of Luneburg's assumptions Heller [2] suggested the investigation of the model

$$\rho(\alpha,\beta) = f(\alpha) - g(\beta) \quad \text{and} \quad \vartheta(\alpha,\beta) = \frac{f(\alpha)+g(\beta)}{2},$$

where $f, g : \left]-\frac{\pi}{2}, \frac{\pi}{2}\right[\to \mathbb{R}$ are strictly increasing and continuous. He also introduced some psychophysical invariances, which are called shift-invariances (two of them are considered in the next paragraph in a slightly different presentation). Applying these invariance properties to Heller's model one can obtain functional equations, which are solved in [1]. It turns out that none of these shift invariances alone but pairs of them force f and g to be continuous affine functions. Thus a pair of these shift invariances and an additional assumption on symmetry reduce Heller's model to Luneburg's one.

In this paper we wish to initiate a generalization of Heller's approach. We do not assume any specific formula for the perceived distance and the perceived direction. We consider these quantities as functions in two variables, satisfying some functional equations (induced by *shift invariances*). However, we do not assume any relation between these two functions, therefore it is natural to consider only one unknown function $F : S \to \mathbb{R}$. Our aim is to prove that any F satisfying the following shift invariance properties can be expressed via a linear function on S. Unfortunately, we need stronger regularity assumptions than those supposed in [1].

2. Functional equations induced by shift invariances

The following concepts are introduced in psychophysics.

DEFINITION 1

$F : S \to \mathbb{R}$ is *α-shift invariant* if

$$F(\alpha,\beta) \leqslant F(\alpha',\beta') \iff F(\alpha+\tau,\beta) \leqslant F(\alpha'+\tau,\beta') \tag{1a}$$

whenever $(\alpha,\beta), (\alpha+\tau,\beta), (\alpha',\beta'), (\alpha'+\tau,\beta') \in S$.

Analogously, $F : S \to \mathbb{R}$ is *β-shift invariant* if

$$F(\alpha,\beta) \leqslant F(\alpha',\beta') \iff F(\alpha,\beta+\tau) \leqslant F(\alpha',\beta'+\tau) \tag{1b}$$

whenever $(\alpha,\beta),(\alpha,\beta+\tau),(\alpha',\beta'),(\alpha',\beta'+\tau)\in S$.

Now we reformulate these properties by means of functional equations.

PROPOSITION 1

$F : S \to \mathbb{R}$ is α-shift invariant if, and only if, there exists a function Ψ_1 in two variables such that the mapping $x \mapsto \Psi_1(x,\tau)$ is strictly increasing for every $\tau \in]-\pi,\pi[$ and

$$F(\alpha+\tau,\beta)=\Psi_1(F(\alpha,\beta),\tau) \tag{2a}$$

holds whenever $(\alpha,\beta),(\alpha+\tau,\beta)\in S$.

Similarly, F is β-shift invariant if, and only if, there exists a function Ψ_2 in two variables such that the mapping $x \mapsto \Psi_2(x,\tau)$ is strictly increasing for every $\tau \in]-\pi,\pi[$ and

$$F(\alpha,\beta+\tau)=\Psi_2(F(\alpha,\beta),\tau) \tag{2b}$$

holds whenever $(\alpha,\beta),(\alpha,\beta+\tau)\in S$.

Proof. Let us assume that $F : S \to \mathbb{R}$ is α-shift invariant. Interchanging the roles of (α,β) and (α',β') in (1a) we obtain

$$F(\alpha,\beta)\geqslant F(\alpha',\beta') \Longleftrightarrow F(\alpha+\tau,\beta)\geqslant F(\alpha'+\tau,\beta') \tag{1a*}$$

on the same domain. The properties (1a) and (1a*) immediately imply

$$F(\alpha,\beta)=F(\alpha',\beta') \Longleftrightarrow F(\alpha+\tau,\beta)=F(\alpha'+\tau,\beta') \tag{3}$$

whenever $(\alpha,\beta),(\alpha+\tau,\beta),(\alpha',\beta'),(\alpha'+\tau,\beta')\in S$. Now for every $\tau\in]-\pi,\pi[$ let

$$X_\tau=\{F(\alpha,\beta)\mid (\alpha,\beta),(\alpha+\tau,\beta)\in S\}$$

and define

$$\Psi_1(x,\tau)=F(\alpha+\tau,\beta)\quad (x=F(\alpha,\beta)\in X_\tau,\ \ \tau\in]-\pi,\pi[).$$

It follows from (3) that the value of $\Psi_1(x,\tau)$ does not depend on the representation $x=F(\alpha,\beta)$ and the mapping $x\mapsto\Psi_1(x,\tau)$ is injective for every $\tau\in]-\pi,\pi[$; while (1a) implies that it is increasing. The definition of Ψ_1 yields (2a). Conversely, if the functional equation (2a) is satisfied by some function Ψ_1 which is strictly increasing in its first variable, we obviously have (1a).

The argument for the β-shift invariant case is the same.

3. Smooth shift-invariant functions

The simplest examples of simultaneously α- and β-shift invariant functions are the restrictions of linear functions to the set S. Moreover, if $F : S \to \mathbb{R}$ is α- and β-shift invariant and $f : F(S) \to \mathbb{R}$ is strictly monotonic, then the composite function $f \circ F$ is α- and β-shift invariant as well. In what follows we prove that under strong (but quite natural) regularity assumptions every α- and β-shift invariant function F can be decomposed in the form $F = f \circ L$, where L is linear and f is strictly monotonic.

THEOREM 1

Let $F : S \to \mathbb{R}$ be continuously differentiable, α- and β-shift invariant, and suppose that $D_1F(\alpha,\beta)D_2F(\alpha,\beta) \neq 0$ for every $(\alpha,\beta) \in S$. Then there exist $a, b \in \mathbb{R}$ and a continuously differentiable, strictly monotonic function f such that

$$F(\alpha,\beta) = f(a\alpha + b\beta) \qquad ((\alpha,\beta) \in S).$$

Proof. Let $I =]-\pi/2, \pi/2[$. Since $D_2F(\alpha,\beta) \neq 0$ for every $(\alpha,\beta) \in S$, the partial derivative $D_2F : S \to \mathbb{R}$ is continuous, and S is connected, the function D_2F has to preserve the sign. Hence either $D_2F(\alpha,\beta) > 0$ for every $(\alpha,\beta) \in S$ or $D_2F(\alpha,\beta) < 0$ for every $(\alpha,\beta) \in S$ (let us note that also D_1F has this property). Therefore F is strictly monotonic in the second variable while keeping the first variable fixed (and vice versa). In particular, for every $x \in F(S)$ and $\alpha \in I$ the equation $F(\alpha,\beta) = x$ is satisfied by at most one $\beta = \phi_x(\alpha)$. As it is familiar from calculus, under our assumptions the domain of the implicit function ϕ_x is an open subset of the interval I, ϕ_x is differentiable, and

$$\phi_x'(\alpha) = -\frac{D_1F(\alpha,\phi_x(\alpha))}{D_2F(\alpha,\phi_x(\alpha))} \tag{4}$$

for every α in the domain of ϕ_x. It follows from (4) that ϕ_x' is continuous and $\phi_x'(\alpha) \neq 0$ for every α, hence the restriction of ϕ_x to any interval is strictly monotonic. Since F is continuous, the graph of ϕ_x, which is actually equal to the set $F^{-1}(x)$, is closed in the subspace topology of S. Thus the domain of ϕ_x is a union of pairwise disjoint open intervals such that for any connected component $]\nu_1,\nu_2[$ we have $(\nu_i, \lim_{\alpha\to\nu_i}\phi_x(\alpha)) \in \partial S$ $(i = 1,2)$, where ∂S denotes the boundary of S and the limit has to be considered only from above or from below, respectively.

The second step is to prove that the functions Ψ_1 and Ψ_2 are continuously differentiable. We present the argument for Ψ_1, using the notation introduced in the proof of Proposition 1. Let $\tau_0 \in]-\pi,\pi[$ and $x_0 \in X_{\tau_0}$ be arbitrary. Due to the definition of X_{τ_0} there exists $(\alpha_0, \beta_0) \in S$ for which $(\alpha_0 + \tau_0, \beta_0) \in S$ and $x_0 = F(\alpha_0, \beta_0)$. Let us define

$$I_0 = \left]\max\{\beta_0 - \alpha_0, \beta_0 - \alpha_0 - \tau_0\}, \min\left\{\frac{\pi}{2} - \alpha_0, \frac{\pi}{2} - \alpha_0 - \tau_0\right\}\right[$$

and $\Phi(\tau) = \Psi_1(x_0, \tau)$ $(\tau \in I_0)$. Then I_0 is an open neighbourhood of 0 and

$$\Phi(\tau) = \Psi_1(x_0, \tau) = \Psi_1(F(\alpha_0, \beta_0), \tau) = F(\alpha_0 + \tau, \beta_0) \in X_{\tau_0} \quad (\tau \in I_0). \quad (5)$$

It follows directly from (5) and from the assumptions on F that the function Φ is continuously differentiable and $\Phi'(\tau) \neq 0$ for every $\tau \in I_0$, hence Φ is strictly monotonic and, in particular, injective. Moreover, its inverse is defined on $\Phi(I_0)$, which is an open neigbourhood of $x_0 = \Phi(0)$, and Φ^{-1} is continuously differentiable as well. For every $x \in \Phi(I_0)$ there exists $\alpha \in I$ such that $\alpha - \alpha_0 \in I_0$ and $x = F(\alpha, \beta_0) = \Phi(\alpha - \alpha_0)$, thus we have

$$\begin{aligned}
\Psi_1(x, \tau) &= \Psi_1(F(\alpha, \beta_0), \tau) \\
&= F(\alpha + \tau, \beta_0) \\
&= F(\alpha_0 + (\alpha - \alpha_0 + \tau), \beta_0) \\
&= \Phi(\alpha - \alpha_0 + \tau) \\
&= \Phi(\Phi^{-1}(x) + \tau)
\end{aligned} \quad (6)$$

with τ chosen arbitrarily from an appropriate neighbourhood of 0. It follows from (6) and from the above listed properties of Φ that the function Ψ_1 has continuous partial derivatives in a neighbourhood of (x_0, τ_0), hence it is also continuously differentiable there. Applying (6) one can easily calculate the partial derivatives of Ψ_1 (and, analogously, those of Ψ_2).

In the third step we prove that $\phi'_x(\alpha)$ does not depend on either x or α. For this purpose let $(\alpha_0, \beta_0), (\alpha_1, \beta_1) \in S$, $\tau = \alpha_1 - \alpha_0$ and $\sigma = \beta_1 - \beta_0$. Furthermore, let $\delta > 0$ be so small that we could define

$$h_1(\alpha) = F(\alpha + \tau, \beta_0 + \sigma) \qquad (\alpha \in]\alpha_0 - \delta, \alpha_0 + \delta[\,)$$

and

$$h_2(\beta) = F(\alpha_0 + \tau, \beta + \sigma) \qquad (\beta \in]\beta_0 - \delta, \beta_0 + \delta[\,).$$

Then

$$h_1(\alpha) = F(\alpha + \tau, \beta_0 + \sigma) = \Psi_1(F(\alpha, \beta_0 + \sigma), \tau) = \Psi_1(\Psi_2(F(\alpha, \beta_0), \sigma), \tau),$$

hence we can calculate

$$
\begin{aligned}
D_1F(\alpha_1,\beta_1)& \\
&= D_1F(\alpha_0+\tau,\beta_0+\sigma)\\
&= h_1'(\alpha_0)\\
&= D_1\Psi_1(\Psi_2(F(\alpha_0,\beta_0),\sigma),\tau)D_1\Psi_2(F(\alpha_0,\beta_0),\sigma)D_1F(\alpha_0,\beta_0).
\end{aligned}
\tag{7}
$$

Similarly,

$$h_2(\beta) = F(\alpha_0+\tau,\beta+\sigma) = \Psi_1(\Psi_2(F(\alpha_0,\beta),\sigma),\tau),$$

hence

$$
\begin{aligned}
D_2F(\alpha_1,\beta_1)& \\
&= D_2F(\alpha_0+\tau,\beta_0+\sigma)\\
&= h_2'(\beta_0)\\
&= D_1\Psi_1(\Psi_2(F(\alpha_0,\beta_0),\sigma),\tau)D_1\Psi_2(F(\alpha_0,\beta_0),\sigma)D_2F(\alpha_0,\beta_0).
\end{aligned}
\tag{8}
$$

Equations (7) and (8) imply

$$\frac{D_1F(\alpha_1,\beta_1)}{D_2F(\alpha_1,\beta_1)} = \frac{D_1F(\alpha_0,\beta_0)}{D_2F(\alpha_0,\beta_0)}, \tag{9}$$

which can be combined with equation (4) to obtain $\phi_x'(\alpha) = c$, where $c \in \mathbb{R}$ does not depend on either x or α.

Now let us consider (α_0,β_0), $(\alpha_1,\beta_1) \in S$ such that $c\alpha_0 - \beta_0 = c\alpha_1 - \beta_1$. We may assume, without loss of generality, that $\alpha_0 \leqslant \alpha_1$. Let $x = F(\alpha_0,\beta_0)$. Let I_x denote the union of the open intervals $]\nu,\mu[$ for which $\alpha_0 \in]\nu,\mu[$ and ϕ_x is defined on $]\nu,\mu[$. Then I_x is an open interval, $\alpha_0 \in I_x$, and ϕ_x is defined on I_x, but the numbers $\inf I_x$ and $\sup I_x$ do not belong to the domain of ϕ_x (in fact, I_x is one component of the domain of ϕ_x, which was described at the end of the first paragraph of the proof). From $\phi_x'(\alpha) = c$ $(\alpha \in I_x)$ we have

$$\phi_x(\alpha) = c\alpha + d_x \qquad (\alpha \in I_x), \tag{10}$$

where d_x is a constant depending on x and I_x. In particular, the definition of x yields $\beta_0 = \phi_x(\alpha_0) = c\alpha_0 + d_x$, whence

$$d_x = \beta_0 - c\alpha_0 = \beta_1 - c\alpha_1.$$

If $\alpha_1 \notin I_x$, then $\alpha_0 < \sup I_x \leqslant \alpha_1$. Let $\xi = \sup I_x$ and $\eta = \lim_{\alpha \to \xi} \phi_x(\alpha)$. Then (10) implies $\eta = c\xi + d_x$, hence (ξ,η) is an element of the segment with extremities (α_0,β_0) and (α_1,β_1). Since S is convex, this yields $(\xi,\eta) \in S$, which contradicts $(\xi,\eta) \in \partial S$ (obtained at the end of the first paragraph of the proof), since S is open. Thus $\alpha_1 \in I_x$. Then (10) yields

$$\phi_x(\alpha_1) = c\alpha_1 + d_x = \beta_1,$$

hence

$$F(\alpha_1, \beta_1) = x = F(\alpha_0, \beta_0).$$

So we have proved that F is compatible with the equivalence relation defined on S by

$$(\alpha_0, \beta_0) \sim (\alpha_1, \beta_1) \Longleftrightarrow \beta_0 - c\alpha_0 = \beta_1 - c\alpha_1.$$

This result can be reformulated in such a way that the formula

$$f(c\alpha - \beta) = F(\alpha, \beta) \quad ((\alpha, \beta) \in S) \tag{11}$$

is a correct definition of the function f. Since S is connected and the linear mapping $(\alpha, \beta) \mapsto c\alpha - \beta$ is continuous, the domain of f is an interval. Moreover, f is continuously differentiable and differentiating (11) with respect to β we obtain

$$f'(c\alpha - \beta) = -D_2F(\alpha, \beta) \neq 0,$$

hence f is strictly monotonic. Thus we have proved the theorem with $a = c$ and $b = -1$.

REMARK 1

Our results remain true when the set S is replaced by any open convex subset of $\mathbb{R}^2$.

PROBLEM 1

It would be interesting to investigate the equations (2a)-(2b) under weaker regularity assumptions. According to author's conjecture the continuous and (in each variable) strictly monotonic solutions should have the same form.

References

[1] J. Aczél, Z. Boros, J. Heller, C.T. Ng, *Functional equations in binocular space perception*. J. Math. Psych. **43** (1999), 71-101.

[2] J. Heller, *On the psychophysics of binocular space perception*, J. Math. Psych., to appear in 1997.

[3] R.K. Luneburg, *Mathematical analysis of binocular vision*, Princeton University Press, Princeton, 1947.

Institute of Mathematics and Informatics
Kossuth Lajos University
H-4010 Debrecen Pf. 12
Hungary
E-mail: boros@math.klte.hu

Manuscript received: November 30, 1997 and in final form: September 18, 1999