Advances in Mathematics 289 (2016) 951–1007

Contents lists available at ScienceDirect

Advances in Mathematics

www.elsevier.com/locate/aim

Operator means of probability measures and generalized Karcher equations

Miklós Pálfia

Department of Mathematics, Kyoto University, Kyoto 606-8502, Japan

A R T I C L E   I N F O

Article history:
Received 18 November 2014
Received in revised form 9 October 2015
Accepted 16 November 2015
Available online 7 December 2015
Communicated by Dan Voiculescu

MSC:
primary 47A64, 46L05
secondary 53C20, 53C35

Keywords:
Operator monotone function
Operator mean
Karcher mean

A B S T R A C T

In this article we consider means of positive bounded linear operators on a Hilbert space. We present a complete theory that provides a framework which extends the theory of the Karcher mean, its approximating matrix power means, and a large part of Kubo–Ando theory to arbitrary many variables, in fact, to the case of probability measures with bounded support on the cone of positive definite operators. This framework characterizes each operator mean extrinsically as unique solutions of generalized Karcher equations which are obtained by exchanging the matrix logarithm function in the Karcher equation to arbitrary operator monotone functions over the positive real half-line. If the underlying Hilbert space is finite dimensional, then these generalized Karcher equations are Riemannian gradients of convex combinations of strictly geodesically convex log-determinant divergence functions, hence these new means are their global minimizers, in analogy to the case of the Karcher mean as pointed out. Our framework is based on fundamental contraction results with respect to the Thompson metric, which provides us nonlinear contraction semigroups in the cone of positive definite operators that form a decreasing net approximating these operator means in the strong topology from above.

E-mail address: palfia.miklos@aut.bme.hu.

http://dx.doi.org/10.1016/j.aim.2015.11.019

1. Introduction

Let E be Hilbert space and $S(E)$ denote the Banach space of bounded linear self-adjoint operators. Let $\mathbb{P} \subseteq S(E)$ denote the cone of positive definite operators on E. In this article we are concerned with means of members of $\mathbb{P}$ that enjoy certain attractive properties that recently became important from the point of view of averaging in the finite dimensional case, see for example [3,13,16,5,4]. Usually the main difficulties here arise from the required property of operator monotonicity i.e., our means must be monotone with respect to the positive definite order on $\mathbb{P}$. The 2-variable theory of such functions is relatively well understood: each such function is represented by an operator monotone function according to the theory of Kubo–Ando [18]. However in the several variable case we have no such characterization of operator monotone functions.

When E is finite dimensional, then there are additional geometrical structures on $\mathbb{P}$ that are used to define n-variable operator means [2,10,22,27]. In this setting $\mathbb{P}$ is just the cone of positive definite n-by-n Hermitian matrices, where n is the dimension of E. It is a smooth manifold as an open subset of the vector space of n-by-n Hermitian matrices (which is just $S(E)$ in this case) and has a Riemannian symmetric space structure $\mathbb{P} \cong \mathrm{GL}(E)/\mathbb{U}$, where $\mathbb{U}$ is the unitary group and $\mathrm{GL}(E)$ is the general linear group over E [8,9]. This symmetric space is nonpositively curved, hence a unique minimizing geodesic between any two points exists [9]. The midpoint operation on this space, which is defined as taking the middle point of the geodesic connecting two points, is the geometric mean of two positive definite matrices [9]. The Riemannian distance on the manifold $\mathbb{P}$ is of the form

$$d(A,B) = \sqrt{Tr \log^2(A^{-1}B)}.$$

Then the (weighted) multivariable geometric mean or Karcher mean of the k-tuple $\mathbb{A} := (A_1, \dots, A_k) \in \mathbb{P}^k$ with respect to the positive probability vector $\omega := (w_1, \dots, w_k)$ is defined as the center of mass

$$\Lambda(w_1, \dots, w_k; A_1, \dots, A_k) = \underset{X \in P(n,\mathbb{C})}{\arg\min} \sum_{i=1}^{k} w_i d^2(X, A_i). \tag{1}$$

The Karcher mean was first considered in this setting in [27,10]. The Karcher mean $\Lambda(\omega; \mathbb{A})$ is also the unique positive definite solution of the corresponding critical point equation called the Karcher equation

$$\sum_{i=1}^{k} w_i \log(X^{-1}A_i) = 0, \tag{2}$$

where the gradient of the function in the minimization problem (1) appears on the left hand side [27]. The (entry-wise) operator monotonicity of this mean in $\mathbb{A}$ with respect to