# Positive Definite Functions on Hilbert Space


B. J. C. Baxter

*Department of Mathematics,*
*Imperial College, London SW7 2BZ,*
*England*

b.baxter@ic.ac.uk

www.ma.ic.ac.uk/~baxter



A function $f : [0, \infty) \to \mathbb{R}$ for which the quadratic form

$$\sum_{j=1}^{n} \sum_{k=1}^{n} a_j a_k f(\|x_j - x_k\|^2)$$

is non-negative, for *any* positive integers $n$ and $d$ and all points $x_1, \ldots, x_n$ lying in $\mathbb{R}^d$, where $\|\cdot\|$ denotes the Euclidean norm, is said to be *positive definite on Hilbert space*. In Schoenberg (1938), it was shown that a function is positive definite on Hilbert space if and only if it is completely monotonic. Unfortunately, Schoenberg's proof was rather complicated. In this paper, we present a short geometric proof of this beautiful fact.


## 1. Introduction

Let $f : [0, \infty) \to \mathbb{R}$ be a function for which the quadratic form

$$\sum_{j=1}^{n} \sum_{k=1}^{n} a_j a_k f(\|x_j - x_k\|^2)$$

is non-negative, for *any* positive integers $n$ and $d$ and all points $x_1, \ldots, x_n$ lying in $\mathbb{R}^d$, where $\|\cdot\|$ denotes the Euclidean norm. In other words, if $H$ is any Hilbert space that is isomorphic to $\ell^2(\mathbb{Z})$, then $f : [0, \infty) \to \mathbb{R}$ has the property that every matrix

$$\Big( f(\|x_j - x_k\|^2) \Big)_{j,k=1}^{n} \tag{1.1}$$

is non-negative definite, for any positive integer $n$ and any points $x_1, \ldots, x_n \in H$. Accordingly, we say that $f$ is *positive definite on Hilbert space*, although we remark that much of the literature prefers the more cumbersome term "positive definite on every $\mathbb{R}^d$". Such functions were characterized by I. J. Schoenberg in a remarkable series of seminal papers collected in Schoenberg (1988), wherein he was able to show that the class of functions positive definite on Hilbert space is precisely the set of *completely monotonic* functions. We recall that $f : (0, \infty) \to \mathbb{R}$ is completely monotonic provided that

$$(-1)^k f^{(k)}(x) \geq 0, \qquad \text{for every } k = 0, 1, \ldots \text{ and for } 0 < x < \infty.$$

The theory of positive definite functions on Hilbert space has had great impact in many fields; see Micchelli (1986) and the survey of metric geometry in Schoenberg (1988).

Unfortunately, Schoenberg's proof was rather complicated. Specifically, he first derived an integral representation for radially symmetric positive definite functions on $\mathbb{R}^d$. A careful limiting argument, similar in flavour to the Central Limit Theorem, was then used to prove that positive definite functions on Hilbert space are Laplace transforms of finite positive Borel measures defined on the half-line $[0, \infty)$. Such Laplace transforms are known to be the completely monotonic functions, by a celebrated theorem of Bernstein (see, for instance, Widder (1946)). In this paper, we present a short direct proof that postive definite functions on Hilbert space are completely monotonic. Further, our proof is based on a geometric construction first used in Baxter (1991).

## 2. Positive Definite Functions on Hilbert Space

In this note, $H$ can be any Hilbert space isomorphic to $\ell^2(\mathbb{Z})$.

**Definition 2.1.** A function $f\colon [0, \infty) \to \mathbb{R}$ will be called *positive definite on Hilbert space* (HPD) if the matrix

$$\Big( f(\|x_j - x_k\|^2) \Big)_{j,k=1}^{n} \tag{2.1}$$

is non-negative definite for every positive integer $n$ and any points $x_1, \ldots, x_n \in H$. We shall call any matrix of the form (2.1) a *distance matrix*.

The classical theory of positive definite functions provides the well-known inequality

$$|f(t)| \leq f(0),$$

but a stronger bound holds for HPD functions.

**Proposition 2.1.** Every HPD function is non-negative.

*Proof.* Let $e_1, e_2, \ldots$ be any orthonormal sequence in $H$ and choose any nonzero real number $\lambda$. Using the relation $f(\|\lambda e_j - \lambda e_k\|^2) = f(2\lambda^2)$, for

$j \neq k$, we find

$$0 \le \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}^T \Big( f(\|\lambda e_j - \lambda e_k\|^2 \Big)_{j,k=1}^n \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} = nf(0) + \frac{1}{2} n(n-1) f(2\lambda^2). \quad (2.2)$$

Hence, for $n \geq 2$, we have

$$f(2\lambda^2) + \frac{2f(0)}{n-1} \geq 0. \quad (2.3)$$

Letting $n \to \infty$, we conclude $f(2\lambda^2) \geq 0$ for every nonzero real number $\lambda$. □

In fact, HPD functions satisfy a *much* stronger property. We recall that, for any $h > 0$, the *forward difference operator* $\Delta_h$ is defined by the equation

$$\Delta_h f(t) := f(t+h) - f(t), \qquad t \geq 0, \quad (2.4)$$

so that $\Delta_h f \colon [0,\infty) \to \mathbb{R}$. Of course, $\Delta_h^m f := \Delta_h(\Delta_h^{m-1} f)$.

**Theorem 2.2.** If $f \colon [0,\infty) \to \mathbb{R}$ is HPD, then $(-1)^m \Delta_h^m f$ is also HPD, for every $h > 0$ and positive integer $m$.

*Proof.* It is sufficient to prove that $-\Delta_h f$ is HPD. To this end, let $x_1, \ldots, x_n$ be any vectors in $H$, and choose any unit vector $y \in H$ that is orthogonal to $x_1, \ldots, x_n$. We now let $A$ denote the $2n \times 2n$ distance matrix generated by the points

$$x_1, \ldots, x_n, x_1 + h^{1/2} y, \ldots, x_n + h^{1/2} y.$$

It is easily checked that

$$A = \begin{pmatrix} B & C \\ C & B \end{pmatrix},$$

where the $n \times n$ matrices $B$ and $C$ are given by the equations

$$B_{jk} = f(\|x_j - x_k\|^2), \qquad 1 \leq j,k \leq n,$$

and

$$C_{jk} = f(\|x_j - (x_k + h^{1/2} y)\|^2) = f(\|x_j - x_k\|^2 + h), \qquad 1 \leq j,k \leq n.$$

Now, given any vector $a \in \mathbb{R}^n$, and since $f$ is HPD, we have

$$0 \leq \begin{pmatrix} a \\ -a \end{pmatrix}^T A \begin{pmatrix} a \\ -a \end{pmatrix} = 2a^T(B - C)a \equiv 2a^T D a,$$

where

$$D_{jk} = -\Delta_h f(\|x_j - x_k\|^2), \qquad 1 \leq j,k \leq n.$$

Hence $-\Delta_h f$ is HPD. □

A standard limiting argument then allows us to deduce that $f$ is completely monotonic. The slick construction of Theorem 2.2 was, in fact, distilled from a more elaborate geometric artifice to be found in Section 3 of Baxter (1991), described below.

We first choose any points $x_1, \ldots, x_n \in H$ and let $a_1, \ldots, a_n$ be real numbers. We now pick a positive integer $m$ and any set of orthonormal vectors $e_1, \ldots, e_m \in H$ that are orthogonal to $x_1, \ldots, x_n$. The vertices of the cube generated by the closed convex hull of $e_1, \ldots, e_m$ will be denoted $y_1, y_2, \ldots, y_{2^m}$, their order being irrelevant. We shall also write

$$y_k = \sum_{\ell=1}^{m} y_k(\ell) e_\ell, \qquad 1 \le k \le 2^m.$$

Let us now introduce the new configuration of points

$$\left\{ x_j + h^{1/2} y_k : 1 \le j \le n, \quad 1 \le k \le 2^m \right\}. \tag{2.5}$$

The coefficient associated with the point $x_j + h^{1/2} y_j$ will be

$$\alpha(j,k) = a_j (-1)^{y_k(1)+\cdots+y_k(m)}, \qquad 1 \le j \le n, \quad 1 \le k \le 2^m. \tag{2.6}$$

Thus

$$0 \le Q := \sum_{j_1=1}^{n} \sum_{k_1=1}^{2^m} \sum_{j_2=1}^{n} \sum_{k_2=1}^{2^m} \alpha(j_1,k_1)\alpha(j_2,k_2) f(\|(x_{j_1}+h^{1/2}y_{k_1})-(x_{j_2}+h^{1/2}y_{k_2})\|^2), \tag{2.7}$$

because $f$ is HPD. However, by construction,

$$\|(x_{j_1} + h^{1/2} y_{k_1}) - (x_{j_2} + h^{1/2} y_{k_2})\|^2 = \|x_{j_1} - x_{j_2}\|^2 + h \|y_{k_1} - y_{k_2}\|^2, \tag{2.8}$$

and the squared distances $\{\|y_{k_1} - y_{k_2}\|^2 : 1 \le k_1, k_2 \le n\}$ take the values $0, 1, 2, \ldots, m$, the distance $\ell$ occuring with frequency $\binom{m}{\ell}$. Further, if $\|y_{k_1} - y_{k_2}\|^2 = \ell$, then

$$(-1)^{y_{k_1}(1)+\cdots+y_{k_1}(m)} (-1)^{y_{k_2}(1)+\cdots+y_{k_2}(m)} = (-1)^\ell,$$

so that

$$\begin{aligned} 0 &\le Q \\ &= \sum_{j_1=1}^{n} \sum_{j_2=1}^{n} a_{j_1} a_{j_2} \sum_{\ell=0}^{m} (-1)^\ell \binom{m}{\ell} f(\|x_{j_1} - x_{j_2}\|^2 + h\ell) \\ &= \sum_{j_1=1}^{n} \sum_{j_2=1}^{n} a_{j_1} a_{j_2} (-1)^m \Delta_h^m f(\|x_{j-1} - x_{j_2}\|^2). \end{aligned} \tag{2.9}$$

Finally, we recall the elementary formula (see, for example, Davis (1975))

$$(-1)^m \Delta_h^m f(t) = \sum_{\ell=0}^{m} (-1)^\ell \binom{m}{\ell} f(t + h\ell).$$

## REFERENCES


B. J. C. Baxter (1991), "Conditionally positive functions and $p$-norm distance matrices", *Constr. Approx.* **7**, 427–440.

P. J. Davis (1975), *Interpolation and Approximation*, Dover.

C. A. Micchelli (1986), "Interpolation of scattered data: distance matrices and conditionally positive functions", *Constr. Approx.* **2**, 11-22.

I. J. Schoenberg (1938), "Metric spaces and completely monotone functions", *Ann. of Math.* **39**, 811-841.

I. J. Schoenberg (1988), *Selected Papers*, Volume 1, ed. C. deBoor, Birkhaüser.

D. V. Widder (1946), *The Laplace Transform*, Princeton University Press.