Annales Univ. Sci. Budapest., Sect. Comp. 46 (2017) 177–190

RATIONAL ZERNIKE FUNCTIONS

Levente Lócsi and Ferenc Schipp
(Budapest, Hungary)

Dedicated to the memory of Professor Antal Iványi

Communicated by Sándor Fridli

(Received May 20, 2017; accepted July 25, 2017)

Abstract. Using the relation between the Poincaré and the Cayley–Klein models of hyperbolic geometry, and the congruence transformations on these models (closely related to Blaschke functions), we present a new hyperbolic implementation of Nelder–Mead simplex method, and construct novel complete orthonormal systems on the disk based on Zernike and Blaschke functions.

1. Introduction

In this paper we construct orthogonal systems on the disk starting from Zernike functions [3, 15, 24]. Blaschke functions play an important role in the construction. These functions can be identified with the congruence transformations on the Poincaré disk model of the Bolyai–Lobachevsky hyperbolic geometry [20, 21]. There is a simple map between the Poincaré model and the Cayley–Klein model. This relation enables us to describe the congruence transformations on the Cayley–Klein model and to implement algorithms

Key words and phrases: Rational Zernike functions, Nelder–Mead algorithm, Blaschke group, hyperbolic geometry.
2010 Mathematics Subject Classification: 30J99, 42C05, 51M10, 65K05.
This research was supported by the Hungarian Scientific Research Funds (OTKA) No. K115804.

with geometric background—such as the Nelder–Mead method—in this model [7, 8, 9, 10, 11, 12, 23].

The discrete Laguerre system defined on the torus is widely used in control theory, and can be obtained by the composition of the power functions and the Blaschke functions [1, 2, 6, 13, 14, 15, 16, 17, 18, 19, 20, 21]. With a similar technique, starting from the Zernike functions and considering the congruence transformations on the Poincaré or Cayley–Klein models, we can construct more general complete orthonormal systems on the disk.

2. The Blaschke group and hyperbolic geometry

Let $\mathbb{C}$ denote the set of complex numbers, $\mathbb{D} := \{\, z \in \mathbb{C} : |z| < 1 \,\}$ the disk, $\overline{\mathbb{D}} := \{\, z \in \mathbb{C} : |z| \leq 1 \,\}$ the closed disk and $\mathbb{T} := \{\, z \in \mathbb{C} : |z| = 1 \,\}$ the torus. The Blaschke functions, defined as

$$B_{\mathfrak{a}}(z) := \epsilon B_a(z),\ B_a(z) := \frac{z-a}{1-\overline{a}z} \quad (z \in \overline{\mathbb{D}}, \mathfrak{a} = (a, \epsilon) \in \mathbb{B} := \mathbb{D} \times \mathbb{T})$$

are bijections on both of the sets $\mathbb{D}$ and $\mathbb{T}$, and the set of functions $\{\, B_{\mathfrak{a}} : \mathfrak{a} \in \mathbb{B} \,\}$ is closed under the operations of forming the composition of functions (denoted by $\circ$) and the inverse. It follows that the sets of restricted functions $\mathfrak{B}_{\mathbb{T}} := \{\, B_{\mathfrak{a}}|\mathbb{T} : \mathfrak{a} \in \mathbb{B} \,\}$ and $\mathfrak{B} := \{\, B_{\mathfrak{a}}|\mathbb{D} : \mathfrak{a} \in \mathbb{B} \,\}$ both form a group with the operation $\circ$. The map $\mathfrak{a} \to B_{\mathfrak{a}}$ induces a group denoted by $(\mathbb{B}, \circ)$ on the set of parameters $\mathbb{B}$ isomorphic to group $(\mathfrak{B}, \circ)$. The group $\mathfrak{B}$ can be considered as the group of congruence transformations in the Poincaré disk model of the Bolyai–Lobachevsky hyperbolic geometry [20]. In this model the hyperbolic lines are the circular segments inside $\mathbb{D}$ intersecting $\mathbb{T}$ perpendicularly, and the diameters of the circle. These can be formalized using functions in $\mathfrak{B}$. Namely the set of all hyperbolic lines is given by

$$\mathfrak{L} := \{\, \ell_{\mathfrak{a}} : \mathfrak{a} \in \mathbb{B} \,\}, \quad \ell_{\mathfrak{a}} := \{\, B_{\mathfrak{a}}(t) : -1 < t < 1 \,\}.$$

The points $\ell_{\mathfrak{a}}(-1) := B_{\mathfrak{a}}(-1)$ and $\ell_{\mathfrak{a}}(1) := B_{\mathfrak{a}}(1) \in \mathbb{T}$ are called the points at infinity of the hyperbolic line $\ell_{\mathfrak{a}}$.

The map $\rho(a, z) := |B_a(z)|$ $(a, z \in \mathbb{D})$ is a metric on the disk, resulting in the complete metric space $(\mathbb{D}, \rho)$, furthermore

$$\rho(B(z_1), B(z_2)) = \rho(z_1, z_2) \quad (z_1, z_2 \in \mathbb{D}, B \in \mathfrak{B}),$$

supporting the claim that the functions in $\mathfrak{B}$ correspond to isometric transformations.

The lines in the Cayley–Klein model are the straight line segments (in the Euclidean sense) connecting two points of $\mathbb{T}$ inside $\mathbb{D}$. It is known that the map

$$G(z) := \frac{2z}{1+|z|^2} \quad (z \in \overline{\mathbb{D}})$$

is a $\overline{\mathbb{D}} \to \overline{\mathbb{D}}$ bijection with the set of fixed points $\mathbb{T} \cup \{0\}$, and its inverse is

$$G^{-1}(z) = \frac{z}{1+\sqrt{1-|z|^2}} \quad (z \in \overline{\mathbb{D}}).$$

The function G maps the hyperbolic lines of the Poincaré model (i.e. the circular arcs intersecting $\mathbb{T}$ perpendicularly) onto the hyperbolic lines of the Cayley–Klein model (i.e. straight line segments connecting points of $\mathbb{T}$). Figure 1 illustrates the map G.

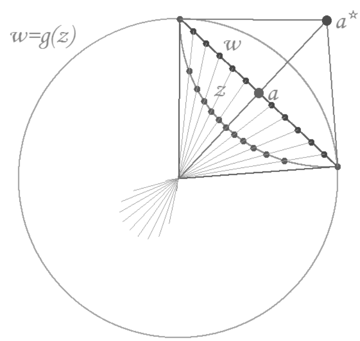

Figure 1. The map G.

The congruence transformations in the Cayley–Klein model can be described by the following functions:

$$\text{(2.1)} \qquad C_{\mathfrak{a}} := G \circ B_{\mathfrak{a}} \circ G^{-1} \quad (\mathfrak{a} \in \mathbb{B}), \quad \mathfrak{C} := \{ C_{\mathfrak{a}} : \mathfrak{a} \in \mathbb{B} \}.$$

The group $(\mathfrak{C}, \circ)$ of these functions is isomorphic to group $(\mathbb{B}, \circ)$, and maps the set of lines of the model onto itself.

Considering a unitary representation of group $(\mathbb{B}, \circ)$, we may construct new orthonormal systems based on an existing orthonormal system in $L^2(\mathbb{D})$. Let $\mathfrak{a}^-$ denote the inverse of $\mathfrak{a} \in \mathbb{B}$, and

$$B'_a(z) := \frac{1-|a|^2}{(1-\overline{a}z)^2} \quad (z \in \mathbb{D})$$

the derivative of B_a. It is easy to see that

$$(2.2)\qquad 1-|B_{\mathfrak{a}}(z)|^2=\frac{(1-|z|^2)(1-|a|^2)}{|1-\bar{a}z|^2}=(1-|z|^2)\,|B'_{\mathfrak{a}}(z)|\quad (z\in\mathbb{D}).$$

We may assign a function $\mathcal{B}\colon \mathbb{R}^2\to\mathbb{R}^2$ to every function $B\colon\mathbb{D}\to\mathbb{D}$ of a complex variable through the map

$$\mathcal{B}(x,y)=(\operatorname{Re}B(z),\operatorname{Im}B(z))\in\mathbb{R}^2\quad (z=x+iy\in\mathbb{D}).$$

If the function $B\colon\mathbb{D}\to\mathbb{D}$ of a complex variable is differentiable, then the function $\mathcal{B}$ of two real variables is also differentiable and—according to the Cauchy–Riemann-equations—for the Jacobi determinant of the derivative $\mathcal{B}'(x,y)$

$$J_B(z):=|\det\mathcal{B}'(x,y)|=|B'(z)|^2\ \ (z=x+iy\in\mathbb{D})$$

holds.

Recall the scalar product on $L^2(\mathbb{D})$:

$$\langle f,g\rangle:=\frac{1}{\pi}\int_{\mathbb{D}}f(z)\,\overline{g}(z)\,\mathrm{d}x\,\mathrm{d}y\quad (z=x+iy,\ f,g\in L^2(\mathbb{D})).$$

Then the operators $U_{\mathfrak{a}}\colon L^2(\mathbb{D})\to L^2(\mathbb{D})$ defined as

$$(2.3)\qquad U_{\mathfrak{a}}f:=B'_{\mathfrak{a}^-}\cdot f\circ B_{\mathfrak{a}^-}\quad (f\in L^2(\mathbb{D}),\ \mathfrak{a}\in\mathbb{B})$$

give a unitary representation of the group $(\mathbb{B},\circ)$ [16, 17, 18, 19], i.e.

$$\begin{aligned}\langle U_{\mathfrak{a}}f,U_{\mathfrak{a}}g\rangle&=\langle f,g\rangle\quad (f,g\in L^2(\mathbb{D}),\ \mathfrak{a}\in\mathbb{B}),\\ U_{\mathfrak{a}_1}(U_{\mathfrak{a}_2}f)&=U_{\mathfrak{a}_1\circ\mathfrak{a}_2}f\quad (f\in L^2(\mathbb{D}),\ \mathfrak{a}_1,\mathfrak{a}_2\in\mathbb{B}).\end{aligned}$$

Indeed, since the absolute value of the Jacobian determinant of $B_{\mathfrak{a}^-}$ considered as $\mathcal{B}_{\mathfrak{a}^-}$ is $\left|B'_{\mathfrak{a}^-}(z)\right|^2$ $(z\in\mathbb{D})$. Thus by integral transformation we have

$$\begin{aligned}\langle f,g\rangle&=\frac{1}{\pi}\int_{\mathbb{D}}f(z)\,\overline{g}(z)\,\mathrm{d}x\,\mathrm{d}y=\frac{1}{\pi}\int_{\mathbb{D}}f(B_{\mathfrak{a}^-}(z))\,\overline{g}(B_{\mathfrak{a}^-}(z))\,\left|B'_{\mathfrak{a}^-}(z)\right|^2\,\mathrm{d}x\,\mathrm{d}y=\\ &=\langle U_{\mathfrak{a}}f,U_{\mathfrak{a}}g\rangle\quad (f,g\in L^2(\mathbb{D}),\ \mathfrak{a}\in\mathbb{B}),\end{aligned}$$

and according to the definition of $U_{\mathfrak{a}}$ and the chain rule

$$\begin{aligned}U_{\mathfrak{a}_1}(U_{\mathfrak{a}_2}f)&=U_{\mathfrak{a}_1}(B'_{\mathfrak{a}_2^-}\cdot f\circ B_{\mathfrak{a}_2^-})=B'_{\mathfrak{a}_1^-}\cdot B'_{\mathfrak{a}_2^-}\circ B_{\mathfrak{a}_1^-}\cdot f\circ B_{\mathfrak{a}_2^-}\circ B_{\mathfrak{a}_1^-}=\\ &=(B_{\mathfrak{a}_2^-}\circ B_{\mathfrak{a}_1^-})'\cdot f\circ B_{\mathfrak{a}_2^-}\circ B_{\mathfrak{a}_1^-}=B'_{(\mathfrak{a}_1\circ\mathfrak{a}_2)^-}\cdot f\circ B_{(\mathfrak{a}_1\circ\mathfrak{a}_2)^-}=\\ &=U_{\mathfrak{a}_1\circ\mathfrak{a}_2}f\quad (f\in L^2(\mathbb{D}),\ \mathfrak{a}_1,\mathfrak{a}_2\in\mathbb{B}).\end{aligned}$$

The weight function

$$\mu(z) := \frac{1}{\left(1-|z|^2\right)^2} \quad (z \in \mathbb{D})$$

satisfies the invariance equation

$$\mu(z) = \mu(B(z))\, J_B(z) \quad (z \in \mathbb{D},\ B \in \mathfrak{B}) \tag{2.4}$$

of the group $(\mathcal{B}, \circ)$ according to (2.2). This implies that the space $L^1_\mu(\mathbb{D})$ with the above μ density function is invariant with respect to the transformations in $\mathfrak{B}$:

$$\int_{\mathbb{D}} f(z)\, \mu(z)\, \mathrm{d}x\, \mathrm{d}y = \int_{\mathbb{D}} f(B(z))\, \mu(z)\, \mathrm{d}x\, \mathrm{d}y \quad (f \in L^1_\mu(\mathbb{D}),\ B \in \mathfrak{B}). \tag{2.5}$$

Indeed, by integral transformation and considering (2.4) we get

$$\int_{\mathbb{D}} f(z)\, \mu(z)\, \mathrm{d}x\, \mathrm{d}y = \int_{\mathbb{D}} f(B(z))\, \mu(B(z))\, J_B(z)\, \mathrm{d}x\, \mathrm{d}y = \int_{\mathbb{D}} f(B(z))\, \mu(z)\, \mathrm{d}x\, \mathrm{d}y.$$

The power functions $h_n(z) := z^n$ $(z \in \mathbb{D},\ n \in \mathbb{N})$ are orthogonal on the space $L^2(\mathbb{D})$. After polar transformation the scalar product has the form

$$\langle f, g\rangle = \frac{1}{\pi} \int_0^{2\pi} \int_0^1 f(re^{i\theta})\, \overline{g}(re^{i\theta})\, r\, \mathrm{d}r\, \mathrm{d}\theta \quad (f, g \in L^2(\mathbb{D})),$$

thus

$$\begin{aligned} \langle h_n, h_m\rangle &= \frac{1}{\pi} \int_0^{2\pi} \int_0^1 r^{n+m} e^{i(n-m)\varphi}\, r\, \mathrm{d}r\, \mathrm{d}\varphi = 0 \quad (n \neq m), \\ \langle h_n, h_n\rangle &= \frac{1}{n+1} \quad (n \in \mathbb{N}). \end{aligned}$$

It follows that the analogue of the discrete Laguerre system on the disk,

$$L_n^{\mathfrak{a}}(z) := (U_{\mathfrak{a}} h_n)(z) = \frac{1-|a|^2}{(1-\overline{a}z)^2} B_a^n(z) \quad (\mathfrak{a} = (-a, 1) \in \mathbb{B},\ z \in \mathbb{D},\ n \in \mathbb{N})$$

is orthogonal in $L^2(\mathbb{D})$. However this system is not complete. In Section 4 we construct a complete orthonormal system in this space starting from Zernike functions.

It follows from the invariance property (2.5) that the maps

$$V_{\mathfrak{a}} f := f \circ B_{\mathfrak{a}^-} \quad (f \in L^2_\mu(\mathbb{D}),\ \mathfrak{a} \in \mathbb{B}) \tag{2.6}$$

form a unitary representation of group $\mathfrak{B}$ in space $L^2_\mu(\mathbb{D})$, i.e.

$$\int_{\mathbb{D}} V_{\mathfrak{a}} f(z)\, \overline{V_{\mathfrak{a}} g(z)}\, \mu(z)\, \mathrm{d}x\, \mathrm{d}y = \int_{\mathbb{D}} f(z) \overline{g(z)}\, \mu(z)\, \mathrm{d}x\, \mathrm{d}y,$$
$$V_{\mathfrak{a}_1}(V_{\mathfrak{a}_2} f) = V_{\mathfrak{a}_1 \circ \mathfrak{a}_2} f \quad (f, g \in L^2_\mu(\mathbb{D}),\, \mathfrak{a}_1, \mathfrak{a}_2 \in \mathbb{B}),$$

and furthermore for the sequence $e_n(z) := (1 - |z|^2) z^n$ $(z \in \mathbb{D},\, n \in \mathbb{N})$

$$\int_{\mathbb{D}} e_n(z) \overline{e}_m(z)\, \mu(z)\, \mathrm{d}x\, \mathrm{d}y = \delta_{nm} \frac{1}{n+1} \quad (m, n \in \mathbb{N}).$$

The functions $G, G^{-1} \colon \mathbb{D} \to \mathbb{D}$ are not differentiable, but at the same time the derivatives of the functions $\mathcal{G}, \mathcal{G}^{-1} \colon \mathbb{R}^2 \to \mathbb{R}^2$ assigned to them exist at every point $z \neq 0, z \in \mathbb{D}$. We will show that the function

$$\nu(v) := \mu(G^{-1}(v))\, J_{G^{-1}}(v) \quad (v \in \mathbb{D}) \tag{2.7}$$

satisfies the invariance equation of group $(\mathfrak{C}, \circ)$, and therefore

$$\int_{\mathbb{D}} g(z)\, \nu(z)\, \mathrm{d}x\, \mathrm{d}y = \int_{\mathbb{D}} g(C(z))\, \nu(z)\, \mathrm{d}x\, \mathrm{d}y \quad (g \in L^1_\nu(\mathbb{D}),\, C \in \mathfrak{C}). \tag{2.8}$$

Indeed, starting from $C = G \circ B \circ G^{-1}$ of (2.1) using the chain rule we get

$$G^{-1} \circ C = B \circ G^{-1}, \quad J_{G^{-1}} \circ C \cdot J_C = J_B \circ G^{-1} \cdot J_{G^{-1}}.$$

Then writing (2.4) for $z = G^{-1}(v)$ and multiplying both sides with $J_{G^{-1}}(v)$:

$$\mu(G^{-1}(v)) \cdot J_{G^{-1}}(v) = \mu(B(G^{-1}(v))) \cdot J_B(G^{-1}(v)) \cdot J_{G^{-1}}(v) =$$
$$= \mu(G^{-1}(C(v))) \cdot J_{G^{-1}}(C(v)) \cdot J_C(v).$$

Thus we have shown that the function ν, indeed, satisfies the equation

$$\nu(v) = \nu(C(v))\, J_C(v) \quad (v \in \mathbb{D},\, C \in \mathfrak{C}).$$

Considering group $(\mathfrak{C}, \circ)$ instead of $(\mathfrak{B}, \circ)$ we may introduce different unitary representations and further orthonormal systems on the disk in a similar manner.

3. The Nelder–Mead algorithm on the hyperbolic plane

The Nelder–Mead simplex algorithm is an unconstrained non-linear optimization method. It was first published in 1965 by Nelder and Mead and since than it has been applied to an enormous amount of optimization problems in the fields of science and engineering [4, 5, 12, 23]. It is also implemented in the `fminsearch` command of numerical computational software such as Matlab. Although practically simple and reliable, this algorithm has very few mathematically proven convergence properties. Some results—for optimization problems of only 1 or 2 variables—appeared in 1998, together with a non-convergent counterexample even for a smooth and convex function [8, 11]. There have been also attempts to modify the behaviour of the original algorithm to enable a wider spectrum of provable properties [23].

The method relies on the comparison of the real function values at the vertices of a non-degenerate simplex in the domain of the function to minimize. (The domain is usually $\mathbb{R}^n$, now we shall consider $\mathbb{D}$ as the hyperbolic plane in the Poincaré and Cayley–Klein models.) The calculations throughout the procedure are usually formalized through linear combinations of the vertices as real vectors, however a pure geometric description is also possible. Below we give a brief sketch of the algorithm using the geometric approach, omitting calculation details.

- We may start with an arbitrary (non-degenerate) simplex, with a hyperbolic triangle on the hyperbolic plane.
- The simplex is updated iteratively replacing the vertex of the worst function value with a better vertex in each step.*
  - The *reflection* of the worst vertex through the centroid of the other vertices is always calculated, and in some cases the update step is completed.
  - In other cases the operations called *expansion, contraction* and *shrink* are applied. Each of these may be geometrically interpreted as finding reflected points or midpoints of line segments.
- The criteria for terminating the iteration is usually that the function values have a small enough standard deviation, or we reach a threshold for the number of steps.

With these steps the simplex 'adapts itself to the local landscape' (defined by the function to minimize) and 'contracts on to the final minimum'. Figure 2 presents an example for some steps of this algorithm on the Euclidean plane optimizing a quadratic function.

*Several vertices are replaced when the shrink operation is applied.

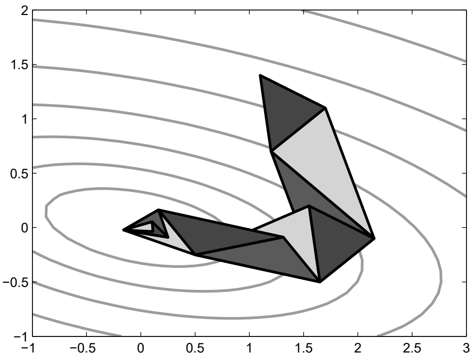

Figure 2. The Nelder–Mead algorithm on the Euclidean plane.

In [10] the Nelder–Mead algorithm adapted to the Poincaré disk model of hyperbolic geometry was presented. Now, applying the maps G and G^{-1} of Section 2, we transform the hyperbolic constructions and the algorithm to the Cayley–Klein model. In our current implementation the constructions rely on the equivalence of these two models: in order to solve a problem in the Cayley–Klein model, we transform the problem to the Poincaré disk model, solve the problem there, and then transform the solution back. The new implementations (Matlab programs) can be downloaded from

`http://numanal.inf.elte.hu/~locsi/hypnm/`.

Figure 3 shows some basic elements of hyperbolic geometry, and one can also examine the Nelder–Mead algorithm operating on the hyperbolic plane (both models) optimizing a function similar to the Rosenbrock function (see [10, 12]). The images present equivalent constellations, i.e. equivalent under the map G. (Not generated by applying the map G to a completed image of course, but invoking the methods implemented on the appropriate model, after finding the equivalent initial settings.)

The operation applied in each step of updating the simplex depends on the locations of the vertices and the function being minimized. There is no prescribed order of applying reflection, expansion, contraction and shrink. Although in case of special initial conditions, when we start the algorithm in equivalent ways, the method makes exactly the same steps in the two models—at least theoretically (numerical errors may occur during practical calculations), and as in the above illustrations. These special equivalent initial conditions for the two hyperbolic variants of the algorithm can be written more precisely as:

- Apply the Nelder–Mead algorithm in the Poincaré disk model with initial simplex $(z_1, z_2, z_3) \in \mathbb{D}^3$ to minimize function $f\colon \mathbb{D} \to \mathbb{R}$.

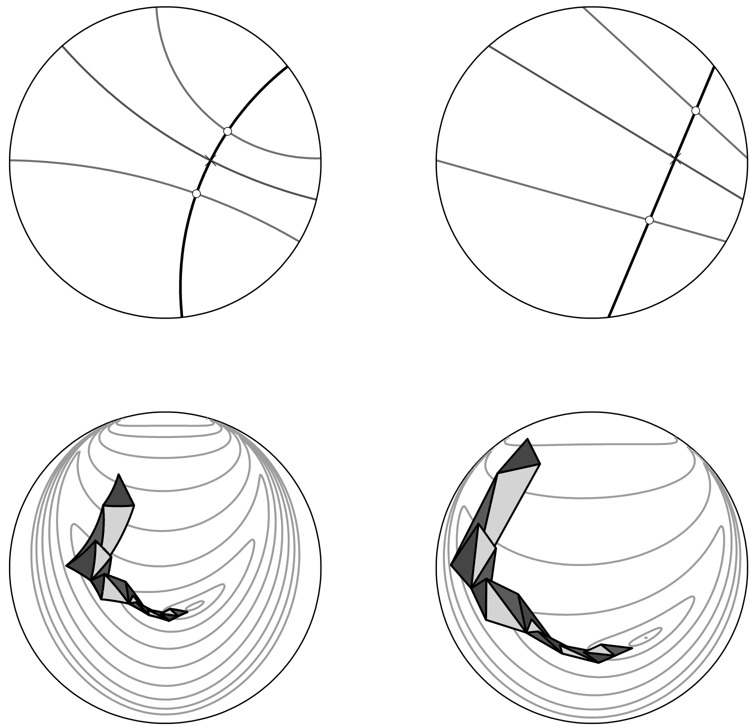

Figure 3. Top row: Some basic elements of hyperbolic geometry. The line fitted on two points, perpendicular lines from the points, and a perpendicular bisector. Bottom row: The progression of the Nelder–Mead method adapted to the hyperbolic plane. Left: Poincaré disk model. Right: Cayley–Klein model.

- Apply the Nelder–Mead algorithm in the Cayley–Klein model with initial simplex $(G(z_1), G(z_2), G(z_3)) \in \mathbb{D}^3$ to minimize function $f \circ G^{-1} : \mathbb{D} \to \mathbb{R}$.

Observe that in both cases the function values at the vertices of the initial simplex evaluate to $f(z_1), f(z_2)$ and $f(z_3)$.

4. Rational Zernike functions

The Zernike functions form an orthonormal system on the disk $\mathbb{D}$ in the space $L^2(\mathbb{D})$. The algebra generated by variables $z \in \overline{\mathbb{D}}$ and $\overline{z} \in \overline{\mathbb{D}}$ shall be denoted by $\mathcal{Z}$:

$$\mathcal{Z} := \operatorname{span}\{\, z^m \overline{z}^n \ :\ m,n \in \mathbb{N} \,\}\,.$$

It is obvious that $\mathcal{Z} \subset L^2(\mathbb{D})$, and $\mathcal{Z}$ is invariant under rotations around the origin. Furthermore the algebra $\mathcal{Z}$ is self-adjoint and separates points of $\overline{\mathbb{D}}$, therefore—according to the Stone–Weierstrass theorem—every continuous function in $C(\overline{\mathbb{D}})$ may be approximated with arbitrary precision with elements of $\mathcal{Z}$.

In 1934 Zernike introduced (see [24]) an orthonormal basis in the algebra $\mathcal{Z}$, the elements of which are called Zernike functions, and are denoted by Y_n^ℓ. Applying the polar transformation $z = re^{i\theta}$, the functions generating $\mathcal{Z}$ can be constructed in the form

$$z^m \overline{z}^n = r^{n+m} e^{i(m-n)\theta} = r^{2n+|\ell|} e^{i\ell\theta} \quad (m,n \in \mathbb{N},\ \ell := m-n \in \mathbb{Z}).$$

Applying the Gram–Schmidt orthogonalization procedure in the space $L_r^2(0,1)$ to the functions

$$r^{|\ell|}, r^{|\ell|+2}, r^{|\ell|+4}, \ldots \quad (0 \le r \le 1),$$

we arrive at the orthonormal set of polynomials $R_n^{|\ell|}$ $(n \in \mathbb{N})$:

$$\int_0^1 R_n^{|\ell|}(r)\, R_m^{|\ell|}(r)\, r\, \mathrm{d}r = \delta_{nm} \quad (n,m \in \mathbb{N}).$$

It follows that the system

$$Y_n^\ell(z) := R_n^{|\ell|}(r)\, e^{i\ell\theta} \quad (z = re^{i\theta} \in \mathbb{D},\ n \in \mathbb{N},\ \ell \in \mathbb{Z})$$

exhibits the below orthogonality property using the polar form of the scalar product of $L^2(\mathbb{D})$:

$$\left\langle Y_n^\ell, Y_m^k \right\rangle = \frac{1}{\pi} \left(\int_0^1 R_n^{|\ell|}(r)\, R_m^{|k|}(r)\, r\, \mathrm{d}r \right) \cdot \left(\int_0^{2\pi} e^{i(\ell-k)\theta}\, \mathrm{d}\theta \right) = 2 \cdot \delta_{n,m}\, \delta_{\ell,k}.$$

From the construction it is clear that $\mathcal{Z} = \operatorname{span}\left\{ Y_n^\ell \ :\ n \in \mathbb{N},\ \ell \in \mathbb{Z} \right\}$, thus the Zernike functions form a closed system in the space $C(\overline{\mathbb{D}})$ and a complete system in the space $L^2(\mathbb{D})$. Then applying the unitary transform (2.3), the systems

$$Z_n^{\ell,\mathfrak{a}} := U_{\mathfrak{a}} Y_n^\ell \quad (\mathfrak{a} \in \mathbb{B},\ n \in \mathbb{N}, \ell \in \mathbb{Z})$$

are again complete orthonormal systems in $L^2(\mathbb{D})$ for every $\mathfrak{a} \in \mathbb{B}$. Similarly in space $L_\mu^2(\mathbb{D})$ starting from the orthonormal system

$$X_n^\ell(z) := (1-|z|^2) Y_n^\ell(z) \quad (z \in \mathbb{D},\ n \in \mathbb{N}, \ell \in \mathbb{Z})$$

and applying the unitary transform (2.6), the resulting system

$$X_n^{\ell,\mathfrak{a}} := V_{\mathfrak{a}} Y_n^{\ell} \quad (\mathfrak{a} \in \mathbb{B},\, n \in \mathbb{N}, \ell \in \mathbb{Z})$$

is orthonormal and complete in the same space. The polynomials $R_n^{|\ell|}$ can also be expressed by means of the Jacobi polynomials $P_n^{\alpha,\beta}$ (see [15, 22]):

$$Y_n^{\ell}(z) = r^{\ell}\, P_n^{0,|\ell|}\,(2r^2-1)\, e^{i\ell\theta} \quad (r = |z|\,,\, \theta = \arg(z),\, n \in \mathbb{N},\, \ell \in \mathbb{Z}).$$

Thus the functions $Z_n^{\ell,\mathfrak{a}}$ can also be written as

$$Z_n^{\ell,\mathfrak{a}}(z) = B'_{\mathfrak{a}}(z)\, r^{\ell}\, P_n^{0,|\ell|}\,(2r^2-1)\, e^{i\ell\theta}$$
$$(r = |B_{\mathfrak{a}}(z)|\,,\, \theta = \arg(B_{\mathfrak{a}}(z)),\, n \in \mathbb{N},\, \ell \in \mathbb{Z}).$$

This concludes the proof of the following

Theorem 1 (Zernike–Poincaré)**.** *For every $\mathfrak{a} \in \mathbb{B}$ the sequences of functions $Z_n^{\ell,\mathfrak{a}}$ $(n \in \mathbb{N},\, \ell \in \mathbb{Z})$ and $X_n^{\ell,\mathfrak{a}}$ $(n \in \mathbb{N},\, \ell \in \mathbb{Z})$ form a complete orthonormal system in spaces $L^2(\mathbb{D})$ and $L^2_{\mu}(\mathbb{D})$ respectively.*

We may construct orthonormal systems in space $L^2_{\nu}(\mathbb{D})$ in a similar manner. Namely, the operators

$$W_{\mathfrak{a}} f := f \circ C_{\mathfrak{a}} \quad (f \in L^2_{\nu}(\mathbb{D}),\, \mathfrak{a} \in \mathbb{B})$$

form a unitary representation of group $(\mathfrak{C}, \circ)$. The functions

$$K_n^{\ell}(z) := \frac{Y_n^{\ell}(z)}{\nu^{1/2}(z)} \quad (z \in \mathbb{D},\, n \in \mathbb{N}, \ell \in \mathbb{Z})$$

are orthonormal in space $L^2_{\nu}(\mathbb{D})$. Considering the above statements, we arrive at the following

Theorem 2 (Zernike–Cayley–Klein)**.** *For every $\mathfrak{a} \in \mathbb{B}$ the sequence of functions $K_n^{\ell,\mathfrak{a}} := W_{\mathfrak{a}} K_n^{\ell}$ $(n \in \mathbb{N},\, \ell \in \mathbb{Z})$ forms a complete orthonormal system in space $L^2_{\nu}(\mathbb{D})$.*

Figure 4 presents two examples of Zernike functions and their unitary transforms.

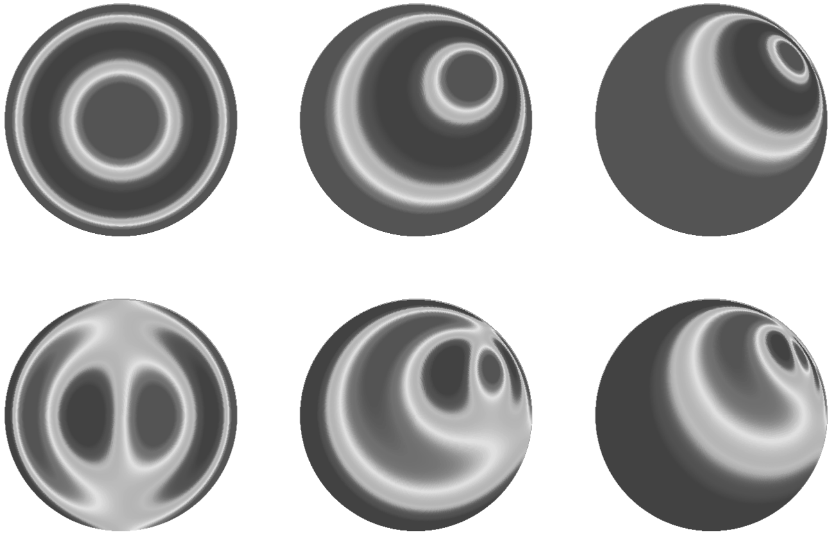

Figure 4. Zernike functions and their unitary transforms as two elements of orthogonal systems. Top row: Y_4^0 and its transforms. Bottom row: Y_5^1 and its transforms. Left: original Zernike function. Middle: Poincaré disk model. Right: Cayley–Klein model.

References

[1] **Bokor, J., F. Schipp and A. Soumelidis,** Pole structure estimation from Laguerre representation using hyperbolic metric on the unit disc, in: IEEE (Ed.) *50th IEEE Conf. on Decision and Control and European Control Conf.* (Orlando, Florida, 2011), IEEE Press, New York, 2011, 2136–2141.

[2] **Eisner, T. and M. Pap,** Discrete orthogonality of the Malmquist–Takenaka system of the upper half plane and rational interpolation, *J. Fourier Anal. Appl.*, **20(1)** (2014), 1–16.

[3] **Fazekas, Z., M. Pap, A. Soumelidis and F. Schipp,** Discrete orthogonality of Zernike functions and its application to corneal measurements. in: S. Ao, L. Gelman (Eds.) *Electronic Engineering and Computing Technology.*, Lect. Notes Electr. Eng. **60**, Springer, Dordrecht, 2010, 455–469.

[4] **Fridli, S., L. Lócsi and F. Schipp,** Rational function systems in ECG processing, in: R. Moreno-Díaz, F. Pichler, A. Quesada-Arencibia (Eds.) *Computer Aided System Theory, EUROCAST 2011* (Las Palmas, Spain, 2011), Springer LNCS **6927**, Springer, Berlin, Heidelberg, 2011, 88–95.

[5] **Fridli, S., P. Kovács, L. Lócsi and F. Schipp,** Rational modeling of multi-lead QRS complexes in ECG signals, *Ann. Univ. Sci. Budapest., Sect. Comp.*, **37** (2012), 145–155.

[6] **Heuberger, P.S.C., P.M.J. Van den Hof and B. Wahlberg (Eds.),** *Modelling and Identification with Rational Orthogonal Basis Functions*, Springer, London, 2005.

[7] **Kovács, P. and L. Lócsi,** RAIT: the Rational Approximation and Interpolation Toolbox for Matlab, with experiments on ECG signals., *Int. J. Adv. in Telecomm. Electrotech. Sig.*, **1(2–3)** (2012), 67–75.

[8] **Lagarias, J.C., J.A. Reeds, M.H. Wright and P.E. Wright,** Convergence properties of the Nelder–Mead algorithm in low dimensions, *SIAM J. Optim.*, **9** (1998), 112–147.

[9] **Lócsi, L.,** Rational FFT implementation in Matlab, *Ann. Univ. Sci. Budapest., Sect. Comp.*, **36** (2012), 241–254.

[10] **Lócsi, L.,** A hyperbolic variant of the Nelder–Mead simplex method in low dimensions *Acta Univ. Sapientiae Math.*, **5(2)** (2013), 169–183.

[11] **McKinnon, K.I.M.,** Convergence of the Nelder–Mead simplex method to a non-stationary point, *SIAM J. Optim.*, **9** (1998), 148–158.

[12] **Nelder, J.A. and R. Mead,** A simplex method for function minimization, *Comput. J.*, **7** (1965), 308–313.

[13] **Pap, M.,** The voice transform generated by a representation of the Blaschke group on the weighted Bergman space., *Ann. Univ. Sci. Budapest., Sect. Comp.*, **33** (2010), 321–342.

[14] **Pap, M.,** Hyperbolic wavelets and multiresolution in $H^2(\mathbb{T})$. *J. Fourier Anal. Appl.*, **17(5)** (2011), 755–776.

[15] **Pap M. and F. Schipp,** Discrete orthogonality of Zernike functions, *Math. Pannon.*, **16(1)** (2005), 137–144.

[16] **Pap, M. and F. Schipp,** The voice transform on the Blaschke group I., *Pure Math. Appl.*, **17(3–4)** (2006), 387–395.

[17] **Pap, M. and F. Schipp,** The voice transform on the Blaschke group II., *Ann. Univ. Sci. Budapest. Sect. Comput.*, **29** (2008), 157–173.

[18] **Pap, M. and F. Schipp,** The voice transform on the Blaschke group III., *Publ. Math. Debrecen*, **75(1–2)** (2009), 263–283.

[19] **Schipp, F.,** Wavelets on the disc, in: K. Hangos, L. Nádai (Eds.) *Proc. Workshop on System and Control Theory (in honor of J. Bokor on his 60th birthday)*, BME AVVC, MTA SZTAKI, Budapest, 2009, 101–109.

[20] **Schipp, F.,** Hyperbolic Wavelets, in: T.M. Rassias, L. Tóth (Eds.) *Topics in Math. Analysis and Applications*, Springer Optim. Appl. **94**, Springer, Berlin, 2014, 633–657.

[21] **Schipp, F.,** Hiperbolikus waveletek, *Alkalmaz. Mat. Lapok*, **32** (2015), 1–39 (in Hungarian).

[22] **Szegő, G.,** *Orthogonal Polynomials*, Amer. Math. Soc. Colloq. Publ. **23**, AMS, Rhode Island, 1975.

[23] **Wright, M.H.,** Nelder, Mead, and the other simplex method, *Doc. Math.*, Extra Volume: Optimization Stories (2012), 271–276.

[24] **Zernike, F.,** Beugungstheorie des Schneidenverfahrens und seiner verbesserten Form, der phasenkontrastmethode, *Physica*, **1(7)** (1934), 689-704.

L. Lócsi and F. Schipp
Department of Numerical Analysis
Eötvös Loránd University
H-1117 Budapest
Pázmány Péter Sétány 1/C
Hungary
locsi@inf.elte.hu
schipp@numanal.inf.elte.hu