Distance-Decreasing Functions on the Hyperbolic Plane

H. S. BEAR

This paper is a study of functions that map the whole hyperbolic plane into itself, and decrease all distances. The important special case consists of the analytic maps of the unit disc into itself.

There is of course a large literature concerning analytic maps of the disc into itself. The kind of result we are interested in started with the work of Julia [9], with subsequent work by Wolff [14] and Carathéodory [6] among others. For an expository article on such maps see Burckel [4].

Consider the following theorem of Wolff.

THEOREM A. *If f is an analytic map of $D=\{z:|z|<1\}$ into D, then either f is conjugate to a rotation, or there is some point $\beta\in\bar{D}=\{z:|z|\le 1\}$ such that the iterates of f converge to β uniformly on compact subsets of D.*

Theorem A and many related results do not depend essentially on the analyticity of f, but only on the fact that f decreases hyperbolic distances in D. Starting with this observation, Beardon [3] recently generalized Theorem A to distance-decreasing functions on Hadamard manifolds; that is, distance-decreasing functions f on (M,d) where (M,d) is a connected, simply connected, complete Riemannian manifold of dimension at least 2 with nonpositive curvature. Goebel and Reich [10] have also treated similar questions in the infinite-dimensional setting.

Our purpose here is to give further properties of distance-decreasing maps of the disc (or hyperbolic plane) into itself. These results will of course apply to all analytic self maps of the disc, and in particular we provide some more details for the result of Theorem A. We take as our axiomatic approach to hyperbolic geometry the metric approach of Birkhoff [4] or Maclane [11]. The basic idea is that each line comes equipped with a linear coordinate system, and there is similarly given a cosmic protractor for measuring angles. The text of Moise [12] gives an excellent comparison of this metric approach with Hilbert's synthetic axioms.

1. The Hyperbolic Plane

Let D be the open unit disc $\{z:|z|<1\}$, which we identify with the hyperbolic plane. The points of the unit circle $\Gamma=\{z:|z|=1\}$ are not points of the

Received September 18, 1990. Final revision received June 17, 1991.
Michigan Math. J. 39 (1992).

hyperbolic plane, but play the role of infinities for the plane which are analogous to $\pm\infty$ on the real line. The hyperbolic lines are arcs of circles orthogonal to Γ. We denote by (α, β) the hyperbolic line (h-line) from $\alpha \in \Gamma$ to $\beta \in \Gamma$.

The Poisson kernel with pole at $\alpha \in \Gamma$ will be denoted $P_\alpha(z)$; thus

$$P_\alpha(z) = \frac{1-|z|^2}{|\alpha - z|^2}. \tag{1}$$

The hyperbolic distance in D is given by

$$d(z, w) = \log \frac{1+\delta(z, w)}{1-\delta(z, w)}, \tag{2}$$

where

$$\delta(z, w) = \frac{|z-w|}{|1-\bar{w}z|}. \tag{3}$$

Equivalently,

$$d(z, w) = 2\tanh^{-1}\delta(z, w). \tag{4}$$

The hyperbolic distance is also given by

$$d(z, w) = \min_{\gamma} \int_\gamma \frac{2|d\zeta|}{1-|\zeta|^2} \tag{5}$$

where the minimum is over all rectifiable curves γ from z to w. The γ that minimizes (5) is the h-segment from z to w. It is shown in [2] that distance is also given by

$$d(z, w) = \max_{\alpha \in \Gamma} \log \frac{P_\alpha(z)}{P_\alpha(w)}. \tag{6}$$

The α that maximizes (6) will be the end of the h-line (α, β) through z and w such that z is closer to α than w is. For all $z \in (\alpha, \beta)$, $P_\alpha(z)P_\beta(z) = \text{const.}$ [2], so $d(z, w)$ is also given by $|\log P_\beta(z)/P_\beta(w)|$.

We assume henceforth that $f: D \to D$ is a distance-decreasing function; that is,

$$d(f(z), f(w)) \leq d(z, w) \tag{7}$$

for all z, w, and *we assume further that f has no fixed points.* If strict inequality holds in (7) we will say that f is *strictly distance-decreasing.* If f is an analytic function on D to D then f is distance-decreasing, so all these results apply in particular to such functions.

Define, for $0 < r < \infty$,

$$R(r) = \max\{d(0, f(z)): d(0, z) \leq r\}. \tag{8}$$

Thus $R(r)$ is the radius of the smallest disc at 0 which contains the image of the r-disc. If $R(r) \leq r$ for some r, then f would be a continuous map of some closed disc into itself and so would have a fixed point, contrary to our assumption. Therefore, for all r,

$$\rho_r = \frac{R(r)}{r} > 1. \tag{9}$$

Since f is distance-decreasing,

$$R(r) \le d(0, f(0)) + r, \tag{10}$$

and

$$R(r+\delta) \le R(r) + \delta. \tag{11}$$

From (9) and (10) we have

$$\rho_r = \frac{R(r)}{r} \to 1 \quad \text{as } r \to \infty. \tag{12}$$

The function ρ_r is of course continuous for $r > 0$, and we show next that ρ_r is strictly decreasing. Let $\Delta > 0$; then

$$\begin{aligned} \rho_{r+\Delta} - \rho_r &= \frac{R(r+\Delta)}{r+\Delta} - \frac{R(r)}{r} \\ &= \frac{r(R(r+\Delta) - R(r)) - R(r)\Delta}{r(r+\Delta)} \\ &\le \frac{r\Delta - R(r)\Delta}{r(r+\Delta)} < 0. \end{aligned} \tag{13}$$

Thus ρ_r is a strictly decreasing continuous function that decreases to 1 as $r \to \infty$.

LEMMA 1. *For each $r > 0$ there is a unique $w_r \in D$ such that*

$$d(0, w_r) \le r \quad \text{and} \quad f(w_r) = \rho_r w_r. \tag{14}$$

The mapping $r \to w_r$ is continuous for $0 < r < \infty$, $|w_r|$ is strictly increasing, and $|w_r| \to 1$ as $r \to \infty$.

Proof. Let $g_r(z) = (r/R(r)) f(z)$, so g_r is a continuous map of the closed hyperbolic r-disc $\{z : d(0,z) \le r\}$ onto itself. By the Brouwer fixed point theorem, g_r has a fixed point, so there is at least one w_r satisfying (14). Suppose there are two points z_1 and z_2 that satisfy (14) for some r: $f(z_1) = \rho_r z_1$ and $f(z_2) = \rho_r z_2$. Since the mapping $z \to tz$ strictly decreases distances if $0 < t < 1$, we have

$$d(z_1, z_2) = d\left(\frac{1}{\rho_r} f(z_1), \frac{1}{\rho_r} f(z_2)\right) < d(f(z_1), f(z_2)). \tag{15}$$

This contradicts our assumption that f is distance-decreasing, so for each r the w_r satisfying (14) is unique.

To show that w_r is a continuous function of r, assume to the contrary that $r_n \to r$ and $|w_{r_n} - w_r| \ge \epsilon$ for all n. We may assume by taking a subsequence that $w_{r_n} \to z \ne w_r$. Then

(16) $$f(w_{r_n}) = \rho_{r_n} w_{r_n} \to \rho_r z = f(z).$$

Since $f(z) = \rho_r z$ and $|z| \le r$, $z = w_r$ and w_r is continuous.

If $|w_r|$ is not strictly increasing, then there are r, s with $0 < s < r$, $\rho_s > \rho_r$, and $|w_s| \ge |w_r|$. It is obvious from hyperbolic trigonometry that $d(tz, t'w) < d(z, w)$ if $0 < t < 1$ and $0 < t' < 1$. Hence

(17) $$d(f(w_r), f(w_s)) = d(\rho_r w_r, \rho_s w_s) > d(w_r, w_s),$$

which is a contradiction. If $|w_r| \not\to 1$ as $r \to \infty$, then there would be an accumulation point in D, and hence a fixed point. □

We will now specialize Beardon's generalization of Theorem A to the plane case. There is a unique point at ∞ for the hyperbolic plane (i.e., a unique point $\beta \in \Gamma$) such that f sends each $z \in D$ toward this β. We will call β the *attractive point* for f. To clarify the sense in which f maps each z toward β, recall ([2], cf. (6) above) that the function $\log P_\beta(z)$ puts a coordinate system on each h-line (α, β) to β. That is, if $z_1, z_2 \in (\alpha, \beta)$ then

(18) $$d(z_1, z_2) = |\log P_\beta(z_2) - \log P_\beta(z_1)|.$$

All the points z with $P_\beta(z) > N$ have coordinates greater than $\log N$ no matter what line (α, β) they lie on. Hence the condition $P_\beta(z) \to \infty$ is a way of saying that z approaches β uniformly on all h-lines to β. We will make the following definition for this kind of hyperbolic limit at infinity, or boundary approach in D.

DEFINITION. For $\{z_n\} \subset D$ and $\beta \in \Gamma$, we say $z_n \to \beta$ hyperbolically if and only if $P_\beta(z_n) \to \infty$.

Notice that hyperbolic convergence as just defined is convergence in a bona fide first countable Hausdorff topology on $D^* = D \cup \Gamma$. A base for the open sets consists of the open sets of D and all sets of the form $\{\beta\} \cup \{z : P_\beta(z) > n\}$ for $\beta \in \Gamma$ and $n = 1, 2, 3, \ldots$. This topology on D^* relativizes to the usual hyperbolic or Euclidean topology on D. We will write

(19) $$\operatorname*{hlim}_{z \to \beta} f(z) = \gamma$$

for this hyperbolic limit; (19) means that $P_\beta(z) \to \infty$ implies $P_\gamma(f(z)) \to \infty$. For $w \in D$,

(20) $$\operatorname*{hlim}_{z \to \beta} f(z) = w$$

means that $P_\beta(z) \to \infty$ implies $d(f(z), w) \to 0$.

In the right half-plane model of the hyperbolic plane, the Poisson kernels are the functions $P_a(x+iy) = x/[x^2 + (y-a)^2]$, $-\infty < a < \infty$, and the function $P_\infty(x+iy) = x$. The horocycle $P_a(z) = n$ is the circle of radius $1/n$ tangent to the imaginary axis at ai. For $\beta = \infty$, $\operatorname{hlim}_{z \to \beta} f(z) = w$ means $f(z) \to w$ as $\operatorname{Re} z \to \infty$.

The hyperbolic approach to a boundary point $\beta \in \Gamma$, or point at infinity, is genuinely intermediate between angular approach and topological approach in the Euclidean metric of $\bar{D}$. There are sequences $\{z_n\}$ which approach $\beta \in \Gamma$ tangentially but not hyperbolically. For example, let $z_n \to \beta$ along a fixed horocycle $P_\beta(z) = k$. There are also sequences that approach $\beta \in \Gamma$ hyperbolically, but not in an angle. For example, let $\beta = 1$ and let z_n be the point where the line through 1 with slope $-n$ intersects the horocycle $P_1(z) = n$. This sequence approaches 1 both tangentially and hyperbolically.

With the above preliminaries, we now state the following.

THEOREM 1 (Beardon). *There is a unique attractive point $\beta \in \Gamma$ such that $P_\beta(f(z)) \geq P_\beta(z)$ for all z, and consequently*

$$\operatorname*{hlim}_{z \to \beta} f(z) = \beta. \tag{21}$$

If f is strictly distance-decreasing, then $P_\beta(f(z)) > P_\beta(z)$ for all z.

Proof. See Ahlfors [1, p. 7]; Beardon [3, p. 146].

COROLLARY 1. *If f is a distance-decreasing map of D into D and f is h-bounded, then f has a fixed point. In particular, an h-bounded entire function on the hyperbolic plane has a fixed point.*

COROLLARY 2 (Liouville's theorem). *A bounded entire function on the complex plane is constant.*

Proof. Assume that f is entire and $|f(z)| \leq M$. Then f has a fixed point in the closed M-disc. Change coordinates so that $f(0) = 0$ and assume without loss that $|f(z)| < 1$ for all z. For all large n, let $g_n(z) = f(nz)$ for $|z| < 1$. Then $g(0) = 0$, and for all z and all $n > |z|$, $d(f(z), 0) = d(g_n(z/n), 0) \leq d(z/n, 0)$. Hence $f(z) \equiv 0$. □

To obtain a second corollary to Theorem 1 we recall the following theorem of Collingwood and Lohwater [8, Thm. 2-22] (see also Rudin [13]). If γ is any arc in D which is tangent to Γ at 1, then there is a $g \in H^\infty$ such that g does not approach a limit along γ or any of its rotations. We will call such a g with $\|g\|_\infty \leq 1$ a Collingwood–Lohwater function for γ.

COROLLARY 3. *If γ is a curve in D that approaches 1 tangentially and hyperbolically, and if g is a Collingwood–Lohwater function for γ, then g has a fixed point.*

Proof. If $\gamma(t)$, $0 \leq t < 1$, is the curve, then the assumption is that $\gamma(t)$ not only approaches Γ tangentially, but also hyperbolically; that is, $P_1(\gamma(t)) \to \infty$ as $t \to 1$. If g does not have a fixed point, then $g(z)$ approaches some $\beta \in \Gamma$ hyperbolically. Rotate γ so that γ approaches β, and hence $P_\beta(\gamma(t)) \to \infty$ as $t \to 1$. Then $g(\gamma(t)) \to \beta$ as $t \to 1$, contrary to our assumption. It follows that g has a fixed point. □

THEOREM 2. *If β is the attractive point for f, then $P_\beta(f(z))/P_\beta(z)$ decreases as z moves toward β along any h-line (α, β). Hence $P_\beta(f(z))/P_\beta(z) \to l_{\alpha\beta} \geq 1$ as $z \to \beta$ along (α, β).*

Proof. Let z_1, z_0 be any two points on (α, β), with z_0 closer to β so that $P_\beta(z_0) > P_\beta(z_1)$. Then

$$\begin{aligned} \log \frac{P_\beta(z_0)}{P_\beta(z_1)} &= d(z_0, z_1) \\ &\geq d(f(z_0), f(z_1)) \\ &\geq \log \frac{P_\beta(f(z_0))}{P_\beta(f(z_1))}. \end{aligned} \tag{22}$$

Hence

$$\frac{P_\beta(z_0)}{P_\beta(z_1)} \geq \frac{P_\beta(f(z_0))}{P_\beta(f(z_1))}, \qquad \frac{P_\beta(f(z_0))}{P_\beta(z_0)} \leq \frac{P_\beta(f(z_1))}{P_\beta(z_1)}. \tag{23}$$

If f is strictly distance-decreasing, then strict inequality holds in (23). Since $P_\beta(f(z))/P_\beta(z) \geq 1$ for all z, the limit $l_{\alpha\beta}$ on (α, β) satisfies $l_{\alpha\beta} \geq 1$. □

Now we return to the curve $w = w_r$, $0 < r < \infty$, of points characterized by $f(w_r) = \rho_r w_r$, $d(0, w_r) \leq r$. We have seen that w_r approaches the unique fixed point β as $r \to \infty$. We now show, taking $\beta = 1$ for convenience, that the curve w_r, $0 < r < \infty$, lies inside the polar curve

$$r = \frac{1 - \sin|\theta|}{\cos \theta}, \tag{24}$$

which goes through $(0, 0)$ tangent to the y-axis and through $(1, 0)$ at angles $\pm\pi/4$ to the x-axis.

THEOREM 3. *If $\beta = 1$ then the curve w_r, $0 < r < \infty$, lies inside the curve* (24), *and hence approaches β within the right angle at β which is split by the radius to β.*

Proof. Assume that $\beta = 1$. Notice that no w_r can lie on the y-axis or to the left of the y-axis. For such points, $\rho_r w_r = f(w_r)$ lies outside the horocycle at 1 through w_r. Let w be any of the w_r, and to be specific assume that w lies in the first quadrant. Let C be the circle centered at 0 through w, and let H be the horocycle at 1 which is perpendicular to C. H cuts C into two arcs, with the smaller arc near 1. If w_r lies on the larger arc, then the point $\rho_r w_r$ will lie outside the horocycle through w_r. Hence w_r must lie on the smaller arcs of the circles at 0 which are cut off by the orthogonal horocycles. If C intersect H at points $re^{i\theta}, re^{-i\theta}$, then an elementary calculation gives (24). The curve (24) is vertical at $(0, 0)$ and has slopes ± 1 at $(1, 0)$. □

The same methods as those used above yield the following extension of Julia's theorem [7, p. 27].

THEOREM 4. *Let f be a distance-decreasing function on D to D with or without fixed points. If there is a sequence $\{z_n\}$ such that $z_n \to \alpha \in \Gamma$, $f(z_n) \to \gamma \in \Gamma$, and* $\liminf(1-|f(z_n)|)/(1-|z_n|) = c < \infty$, *then* $c > 0$ *and* $\operatorname{hlim}_{z\to\alpha} f(z) = \gamma$.

Proof. Assume $\{z_n\}$ is as above. For any distance-decreasing function, we have [7, p. 25]

(25) $$\frac{1-|f(z)|}{1-|z|} \geq \frac{1-|f(0)|}{1+|f(0)|}.$$

Hence $c \geq (1-|f(0)|)/(1+|f(0)|) > 0$.

For any real sequence $\{s_n\}$ and any k with $0 < k < \infty$,

(26) $$\frac{1-s_n^2}{1-|z_n|^2} \to k \quad \text{if and only if} \quad \frac{1-s_n}{1-|z_n|} \to k.$$

Hence, if the limits (26) hold,

(27) $$\limsup_{n\to\infty} \frac{1-s_n^2}{1-|f(z_n)|^2} = \limsup_{n\to\infty} \frac{1-s_n^2}{1-|z_n|^2} \cdot \frac{1-|z_n|^2}{1-|f(z_n)|^2} = \frac{k}{c}.$$

Now recall that

(28) $$\frac{|z-z_n|}{|1-\bar{z}_n z_n|} < s_n \quad \text{if and only if} \quad \frac{1-|z|^2}{|1-\bar{z}_n z|^2} > \frac{1-s_n^2}{1-|z_n|^2}.$$

Pick a sequence $\{s_n\}$ so that the right side above equals k for all n. If $P_\alpha(z) > k$, then the right inequality of (28) holds for all large n, and consequently the left inequality holds for all large n, and even so with $f(z_n)$ and $f(z)$ in place of z_n and z because f is distance-decreasing. Taking the limit in the right inequality of (28) with $f(z)$ and $f(z_n)$ in place of z and z_n gives, using (27), $P_\gamma(f(z)) \geq k/c$. Thus $P_\alpha(z) \geq k$ implies $P_\gamma(f(z)) \geq k/c$. Hence

(29) $$\operatorname{hlim}_{z\to\alpha} f(z) = \gamma,$$

and furthermore

(30) $$\frac{P_\gamma(f(z))}{P_\alpha(z)} \geq \frac{1}{c}$$

for all z. The same argument used in the proof of Theorem 1 shows that strict inequality holds in (30) if f is strictly decreasing. ❐

THEOREM 5. *Let f be a distance-decreasing function on D to D with or without fixed points. If there is one sequence $\{z_n\}$ which approaches α within an angle such that*

(31) $$\frac{P_\gamma(f(z_n))}{P_\alpha(z_n)} \geq c > 0,$$

then $\operatorname{hlim}_{z\to\alpha} f(z)=\gamma$. *That is, condition* (31) *for one angular approach sequence insures that the limit is uniform along all h-lines to* α.

Proof. By Theorem 4 it suffices to show that $\liminf(1-|f(z_n)|)/(1-|z_n|)<\infty$. We take $\alpha=1$ and assume that $z_n\to 1$ within an angle whose sides make an angle θ_0 with the real axis. Let z_n lie on the horocycle $P_1(z)=k_n>1$ and let θ_n be the angle between the real axis and the line from 1 to z_n, so $\theta_n\le\theta_0$. Let x_n, $0<x_n<1$, be the point on the horocycle $P_1(z)=k_n$ which is closest to 0. Then $(1+x_n)/(1-x_n)=k_n$. By the triangle inequality in the triangle $0, x_n, z_n$ we have

$$x_n+(1-x_n)\sin\theta_n\ge |z_n|$$

$$\begin{aligned} 1-|z_n| &\ge (1-x_n)(1-\sin\theta_n) \\ &\ge (1-x_n)(1-\sin\theta_0). \end{aligned} \tag{32}$$

By hypothesis, $P_\gamma(f(z_n))\ge cP_1(z_n)=ck_n$, so $f(z_n)$ lies inside the horocycle $P_\gamma(z)=ck_n$. If y_n is the modulus of the closest point of this horocycle to 0, then

$$\frac{1+y_n}{1-y_n}=ck_n=c\frac{(1+x_n)}{1-x_n},$$

$$\frac{1-y_n}{1-x_n}=\frac{1}{c}\frac{1+y_n}{1+x_n}. \tag{33}$$

Since $f(z_n)$ lies inside the horocycle whose farthest point from γ is y_n,

$$1-|f(z_n)|\le 1-y_n,$$

and hence

$$\begin{aligned} \frac{1-|f(z_n)|}{1-|z_n|} &\le \frac{1-y_n}{(1-x_n)(1-\sin\theta_0)} \\ &= \frac{1}{c}\left(\frac{1+y_n}{1+x_n}\right)\frac{1}{1-\sin\theta_0}. \end{aligned} \tag{34}$$

Hence $(1-|f(z_n)|)/(1-|z_n|)$ is bounded, and the result follows from Theorem 4. □

References

1. L. V. Ahlfors, *Conformal invariants; topics in geometric function theory,* McGraw-Hill, New York, 1973.
2. H. S. Bear, *Part metric and hyperbolic metric,* Amer. Math. Monthly (to appear).
3. A. F. Beardon, *Iteration of contractions and analytic maps,* J. London Math. Soc. (2) 41 (1990), 141–150.
4. G. D. Birkhoff, *A set of postulates for plane geometry based on scale and protractor,* Ann. of Math. (2) 33 (1932), 329–345.

5. R. B. Burckel, *Iterating analytic self-maps of discs,* Amer. Math. Monthly 88 (1981), 396–407.
6. C. Carathéodory, *Über die Winkeldervierten von beschränkten analytischen Funktionen,* Sitzsber. Preus. Akad. Wiss., Physikmath., Kl, 1929, 39–54.
7. ———, *Theory of functions,* v. 2, Chelsea, New York, 1954.
8. E. F. Collingwood and A. J. Lohwater, *The theory of cluster sets,* Cambridge Univ. Press, Cambridge, 1966.
9. G. Julia, *Extension nouvelle d'un lemme de Schwarz,* Acta Math. 42 (1920), 349–355.
10. K. Goebel and S. Reich, *Uniform convexity, hyperbolic geometry, and non-expansive mappings,* Marcel Dekker, New York, 1984.
11. S. MacLane, *Metric postulates for plane geometry,* Amer. Math. Monthly 66 (1959), 543–555.
12. E. E. Moise, *Elementary geometry from an advanced standpoint,* 2nd ed., Addison-Wesley, Reading, MA, 1974.
13. W. Rudin, *Inner function images of radii,* Math. Proc. Cambridge Philos. Soc. 85 (1979), 357–360.
14. J. Wolff, *Sur l'itération des fonctions holomorphes dans une région, et dont les valeurs appartiennent à cette région,* C. R. Acad. Sci. Paris Sér. I Math. 182 (1926), 42–43.

Department of Mathematics
University of Hawaii
Honolulu, Hawaii 96822