A geometric interpretation of the complex dilatation

Nikolai V. Ivanov

1. Introduction

Complex analytic notions sometimes turn out to be a stumbling block for the geometrically minded mathematicians. One of such notions is that of Schwarzian derivative, the geometric meaning of which was clarified by J.H. Hubbard [H] and W.P. Thurston [T]. The goal of this note is to provide a geometric interpretation of the *complex dilatation*.

The usual geometric discussion of the dilatation of a map $f\colon U \longrightarrow \mathbf{C}$ from an open subset U of $\mathbf{C}$ to $\mathbf{C}$ in terms of the images of infinitesimal circles succeeds only partially. It does not lead naturally to the complex dilatation

(1) $\quad \mu_f = \overline{\partial} f/\partial f,$

which serves as the main technical tool.

One can define the *conformal dilatation* of a smooth map $f\colon M \longrightarrow R$ at a point $z \in M$ from a smooth two-dimensional manifold M to a Riemann surface R as a point in a canonically defined hyperbolic plane H_z associated to the tangent space to M at z. Almost by the definition, this conformal dilatation is closely related to the distortion of infinitesimal circles. If M is an open subset in $\mathbf{C}$, then the hyperbolic plane H_z can be canonically identified with the unit disc in $\mathbf{C}$, considered as the conformal disc model of the hyperbolic plane. In this case our conformal dilatation turns out to be a complex number and deserves to be called the *complex dilatation*. The discussion in this note is largely limited by the case when M is an open subset in $\mathbf{C}$ and $R = \mathbf{C}$.

It turns out that the complex dilatation defined in this way is related to the distortion of infinitesimal circles in exactly the same way as the classical one. It follows that our

 Neither the work reported in this paper, nor its preparation were supported by any governmental or non-governmental agency, foundation, or institution.

definition is equivalent to the classical one. While the classical definition (1) is very efficient from the analytical point of view, the above approach reveals the geometric structure behind it.

I will not be surprised if the ideas suggested below are well known in some quarters, but I am not aware of any presentation of them.* The purpose of this note is not to establish priority, but to make these ideas available. I pondered on these matters for a while, but decided to put this material in a more definite form only after J.D. McCarthy rekindled my interest in it during the Autumn of 1995. I am very grateful to him for several stimulating discussions of the complex dilatation and related topics.

2. Conformal structures on a vector space

2.1. Conformal structures and the discriminant. Let us begin by recalling some basic notions related to conformal structures. By the definition, a *conformal structure* on a real vector space V is a positive definite real quadratic form on V considered up to multiplication by a positive real number. Equivalently, a conformal structure on V is a definite real quadratic form on V considered up to multiplication by a non-zero real number. Therefore the set of all conformal structures on V is a subset of the projective space $PQ(V)$ associated to the vector space $Q(V)$ of quadratic forms on V. Namely, it is equal to the image in $PQ(V)$ of the set of all definite forms. We will denote the set of all conformal structures on V by $C(V)$.

It is well known that $C(V)$ is a component of the complement in $PQ(V)$ of the so-called *discriminant variety* $\mathcal{D}(V)$ defined by the homogeneous equation $d(q) = 0$, where

$$d\colon Q(V) \to \mathbf{R}$$

is the *determinant function*, defined as follows (see [Ca], Section 1.2). Let us fix a basis $v_1, v_2, \ldots, v_n$ of V. A quadratic form $q \in Q(V)$ leads to a bilinear symmetric form

$$P_q(u, v) = \big(q(u+v) - q(u) - q(v)\big)/2,$$

called the *polarization* of q. By the definition,

$$d(q) = \det\big(P_q(v_i, v_j)\big).$$

It is easy to see that a change of the basis of V changes the determinant function by multiplying it by λ^2, where λ is the determinant of the matrix expressing one basis in

* This was written at the end of 1995 and is still true at the end of 2016. (Added November 6, 2016.)

terms of the other. In particular, $\mathcal{D}(V)$ does not depend on the choice of a basis.

2.2. The determinant in the two-dimensional case. For the rest of the paper, we will assume that V is a two-dimensional real vector space. A crucial fact is that in this case *the determinant function is itself a quadratic form on* $Q(V)$, as we will see in a moment.

It is sufficient to consider the case $V = \mathbf{R}^2$. In this case one can write down an explicit formula for the discriminant. To begin with, one can write any quadratic form q on $\mathbf{R}^2$ in the form $q = aX^2 + bXY + cY^2$, where X, Y are the standard coordinate functions on $\mathbf{R}^2$. Then the matrix of the bilinear form P_q in the standard basis of $\mathbf{R}^2$ is

$$\begin{pmatrix} a & b/2 \\ b/2 & c \end{pmatrix}$$

and its determinant is equal to $ac - b^2/4$. Clearly, (a, b, c) can be used as coordinates on $Q(\mathbf{R}^2)$ and in these coordinates the determinant with respect to the standard basis of $\mathbf{R}^2$ has the form

$$d(a, b, c) = ac - b^2/4.$$

It will be more convenient to deal with the *discriminant*

$$D = -4\,d$$

instead of the determinant (cf. J.W.S. Cassels [Ca], Section 14.1). Clearly,

$$D(a, b, c) = b^2 - 4ac$$

and the change of variables

$$(a, b, c) = \big((t + x)/2,\ y,\ (t - x)/2\big) \tag{2}$$

puts D into the diagonal form

$$D(x, y, t) = x^2 + y^2 - t^2.$$

In particular, the signature of the quadratic form D is $(+, +, -)$.

The set of all definite (positive or negative) quadratic forms q on V is described by the inequality $D(q) < 0$ or, in the coordinates (a, b, c), by the familiar inequality

$$b^2 - 4ac < 0.$$

In the coordinates (x, y, t) it takes the form $x^2 + y^2 - t^2 < 0$. In particular, in the coordinates (x, y, t) we have

$$C(V) = \{ [x : y : t] \in PQ(V) \mid x^2 + y^2 - t^2 < 0 \}.$$

2.3. The conformal structures in the two-dimensional case. If V is two-dimensional, then the equation $d(q) = 0$, or, what is the same, the equation $D(q) = 0$ is a quadratic equation. In other words, the discriminant variety $\mathcal{D}(V)$ i s a conic in the projective plane $PQ(V)$, and the set $C(V)$ of conformal structures on V is a component of the complement $PQ(V) \setminus \mathcal{D}(V)$ of this conic.

More precisely, $C(V)$ is the set of lines $L \in PQ(V)$ such that $D(q) < 0$ for every $q \in L \setminus \{0\}$, because the set of definite forms q is described by $D(q) < 0$. Since D is a quadratic form of the signature $(+, +, -)$, the component $C(V)$ of $PQ(V) \setminus \mathcal{D}(V)$ is naturally a *model of the hyperbolic plane*.

See M. Berger [B], Section 19.2, or, for a discussion in the context of elementary algebraic geometry, H. Clemens [Cl], Sections 1.7 and 1.8. It is worth to point out the the hyperbolic distance is defined in terms of the quadratic form only without recourse to the coordinates used above preceding discussion, cf. [B], Proposition 19.2.5.

In the case of $V = \mathbf{R}^2$ one can use the coordinates (x, y, t) from 2.2 to identify this model with the projective disc model of the hyperbolic plane. See [B], Section 19.2 for the latter. Namely, if $D(x, y, t) < 0$, then $t \neq 0$ and

$$(x/t)^2 + (y/t)^2 < 1.$$

It follows that any line $L \in C(\mathbf{R}^2)$ intersects the plane

$$H = \{ (x, y, t) \mid t = 1 \}$$

in exactly one point and this point belongs to the unit disc with the center $(0, 0, 1)$ in this plane. Hence we can identify $C(\mathbf{R}^2)$ with this disc. Forgetting the last coordinate 1, we can identify H with $\mathbf{R}^2$, and then $C(\mathbf{R}^2)$ turns out to be identified with the unit disc with the center 0 in $\mathbf{R}^2$. If we consider this unit disc as the projective disc model, this identification respects the structure of a hyperbolic plane, i.e. it identifies lines with lines and preserves the hyperbolic distance. Cf. [B], Section 19.2. In the coordinates (a, b, c) on $Q(\mathbf{R}^2)$ the resulting map from $C(\mathbf{R}^2)$ to the unit disc in $\mathbf{R}^2$ is given by the formula

$$(3) \quad [a : b : c] \longmapsto \left(\frac{a - c}{a + c}, \ \frac{b}{a + c} \right)$$

as it directly follows from (2) and the previous paragraph.

3. From the projective to the conformal disc model

3.1. Quadratic forms on $\mathbf{R}^2$. It is well known that any quadratic form on $\mathbf{R}^2$ can be put into a diagonal form by a rotation. Hence any quadratic form on $\mathbf{R}^2$ can be presented in the standard coordinates X, Y in the form

$$A^2(X\cos\theta + Y\sin\theta)^2 + B^2(-X\sin\theta + Y\cos\theta)^2 \tag{4}$$

for some $A, B > 0$ and some angle θ.

The quadratic form (4) defines a conformal structure on $\mathbf{R}^2$ and, in view of the subsection 2.3, a point in the projective disc model of the hyperbolic plane. The goal of this section is to find the corresponding point in the conformal disc model of the hyperbolic plane. This will be done in Lemma 3.3.

3.2. Lemma. *If (r, θ) and (R, Θ) are polar coordinates of two points in the conformal disc and the projective disc model respectively corresponding one to the other under the standard isomorphism between these models, then*

$$\Theta = \theta \quad \textit{and} \quad R = \frac{2r}{1 + r^2}\,. \tag{5}$$

Proof. Let us recall the standard isomorphism between the conformal and the projective disc models (cf. [B], Section 19.6.1). Let us dentify $\mathbf{R}^2$ with the plane

$$\{(x, y, z) \mid z = 0\}$$

in $\mathbf{R}^3$ and let

$$S^2_- = \{(x, y, z) \mid x^2 + y^2 + z^2 = 1,\ z < 0\}$$

be the lower unit hemisphere in $\mathbf{R}^3$. Let

$$\pi : (x, y, z) \longmapsto (x, y)$$

be the orthogonal projection from S^2_- onto the unit disc in $\mathbf{R}^2$, and let f be the radial projection with the center at $(0, 0, 1)$ from S^2_- onto the same unit disc. Then the map $\Omega = \pi \circ f^{-1}$ from the unit disc to itself, considered as a map from the conformal model to the projective one, establishes the standard isomorphism between them.

Therefore, we need to prove that $\Omega(r, \theta) = (R, \Theta)$ if (r, θ) and (R, Θ) are related by (5). Equivalently, we need to prove that $f^{-1}(r, \theta) = \pi^{-1}(R, \Theta)$ under the same assumption. It is sufficient to show that the line through $(0, 0, 1)$ and (r, θ)

intersects the vertical line through (R, Θ) in a point on the lower hemisphere S^2_-. Since $\theta = \Theta$, both lines are contained in the same vertical plane, and we can restrict our attention to this plane.

Therefore, it is sufficient to prove that on the standard coordinate plane, the point of the intersection of the line through $(0, 1)$ and $(r, 0)$ and the vertical line through

$$\left(\frac{2r}{1 + r^2}, \ 0\right)$$

lies the lower unit half-circle. The points on the second line have the form

$$\left(\frac{2r}{1 + r^2}, \ z\right).$$

Such a point lies on the first line if and only if

$$\frac{2r}{1 + r^2}\Big/(1 - z) = r$$

i.e. if and only if $\dfrac{2}{1 + r^2} = 1 - z$ or, equivalently,

$$z = \frac{r^2 - 1}{r^2 + 1}.$$

It follows that the intersection point of two lines is

$$\left(\frac{2r}{1 + r^2}, \ \frac{r^2 - 1}{r^2 + 1}\right).$$

A well known calculation shows that it is contained in the unit circle, and since

$$\frac{r^2 - 1}{r^2 + 1} < 0$$

for $r < 1$, it is contained in its lower half-circle. The lemma follows. ■

3.3. Lemma. *The point in the conformal disc model of the hyperbolic plane corresponding to the conformal structure on* $\mathbf{R}^2$ *defined by the quadratic form* (4) *is equal to*

$$\frac{\gamma - 1}{\gamma + 1}(\cos 2\theta, \ \sin 2\theta),$$

where $\gamma = A/B$.

Proof. The quadratic form (4) is equal to

$$(A^2\cos^2\theta + B\sin^2\theta)X^2 + (2A^2\cos\theta\sin\theta - 2B^2\cos\theta\sin\theta)XY$$

$$+ (A^2\sin^2\theta + B^2\cos^2\theta)Y^2$$

$$= (A^2\cos^2\theta + B^2\sin^2\theta)X^2 + (A^2 - B^2)\sin 2\theta XY + (A^2\sin^2\theta + B^2\cos^2\theta)Y^2.$$

Using (3), we see that the point in the projective disc model corresponding to the conformal structure defined by (4) is equal to

$$\left(\frac{(A^2 - B^2)(\cos^2\theta - \sin^2\theta)}{A^2 + B^2}, \ \frac{(A^2 - B^2)\sin 2\theta}{A^2 + B^2}\right)$$

$$= \frac{A^2 - B^2}{A^2 + B^2}(\cos 2\theta, \ \sin 2\theta).$$

By Lemma 3.2, the point in the conformal disc model corresponding to the conformal structure defined by (4) is equal to

(6) $\quad r(\cos 2\theta, \ \sin 2\theta),$

where r is determined by

$$\frac{A^2 - B^2}{A^2 + B^2} = \frac{2r}{1 + r^2},$$

or, in terms of $\gamma = A/B$,

$$\frac{\gamma^2 - 1}{\gamma^2 + 1} = \frac{2r}{1 + r^2}.$$

Solving the last equation for γ^2, we get

$$\gamma^2 = \frac{(1 + r)^2}{(1 - r)^2}.$$

After taking into account that γ is positive together with A, B, we see that

$$\gamma = \frac{1 + r}{1 - r}.$$

The last equation is equivalent to $r = \dfrac{\gamma - 1}{\gamma + 1}$. Together with (6) this completes the proof of lemma. ∎

4. Dilatation

4.1. Conformal and complex dilatation of a linear map. Let V, W be two real vector spaces of dimension 2. Suppose that W is equipped with a conformal structure. Let $q_W \in Q(W)$ be a quadratic form defining this conformal structure. If $F \colon V \longrightarrow W$ is an isomorphism of real vector spaces, then the induced by F from q_W quadratic form

$$F^* q_W : v \longmapsto q_W(F(v))$$

defines a conformal structure on V. As we saw in Section 2, this conformal structure gives rise to a point in the space $C(V)$, which has a natural structure of a model of hyperbolic plane. We call this point the *conformal dilatation* of f.

In the case of $V = \mathbf{R}^2$ this model of hyperbolic plane can be identified with the projective disc model as in the subsection 2.3. By using the standard identification of the projective disc model with the conformal disc model we can identify this hyperbolic plane also with the latter model. Therefore the conformal structure $F^* q_W$ gives rise to a point in each of these two models.

Let us define the *complex dilatation* μ_F of F as the corresponding point in the conformal disc model. It is a point in the open unit disc in $\mathbf{R}^2$ and, if we identify $\mathbf{R}^2$ with $\mathbf{C}$ in the usual manner, it can be considered as a complex number, whence the name. If F is not an isomorphism, its dilatation is not defined.

4.2. The general case. With the complex analysis and the theory of quasiconformal maps in mind, it is natural to assume that both V and W are complex vector spaces of complex dimension 1 equipped with their standard conformal structures and that $F \colon V \longrightarrow W$ is a real linear map. In this case the conformal dilatation of F belongs to the hyperbolic plane $C(V)$ with *a distinguished point*, namely, with the point corresponding to the standard conformal structure on V. The distinguished point allows to define another structure of a model of hyperbolic plane on $C(V)$, similar to the standard conformal disc model, and an analogue of the standard isomorphism between the projective disc model and the conformal disc model. The image μ_F of the conformal dilatation of F under this isomorphism is a generalization of the complex dilatation. In general, μ_F is not well defined as a complex number, but $|\mu_F|$ is still well defined.†

4.3. Dilatation and the distortion of circles. Now we will relate the complex dilatation μ_F of a real linear map $F \colon V \longrightarrow W$ to the distortion of circles by F. We will assume that a quadratic form $q_W \in Q(W)$ is given and that $V = \mathbf{R}^2$. The form q_W defines a conformal structure and a structure of a Euclidean plane on W.

† In fact, in the general case μ_F is well defined as a tensor, as it should be. (Added November 6, 2016.)

It is well know that F maps the unit circle in $\mathbf{R}^2$ with the center 0 into an ellipse. Let θ be the angle between the positive X-half-axis in $\mathbf{R}^2$ and the direction mapped by F to the direction of the major axis of this ellipse. This angle is determined up to adding a multiple of π and $\theta + (\pi/2)$ is the similar angle related to the minor axis. Moreover,

$$\begin{aligned} F^* q_W &= A^2 \big(X \cos\theta + Y \sin\theta\big)^2 + B^2 \big(X \cos(\theta + (\pi/2)) + Y \sin(\theta + (\pi/2)\big)^2 \\ &= A^2 \big(X \cos\theta + Y \sin\theta\big)^2 + B^2 \big(-X \sin\theta + Y \cos\theta\big)^2, \end{aligned}$$

where A is the length of the major half-axis with respect to q_W and B is the length of the minor half-axis. Lemma 3.3 implies that

$$(7) \qquad \mu_F = \frac{\gamma - 1}{\gamma + 1} (\cos 2\theta, \ \sin 2\theta),$$

where $\gamma = A/B$. Since F maps all other circles with center 0 into homothetic ellipses, μ_F can be determined in the same manner by the action of F on any such a circle.

4.4. Dilatation of smooth maps $\mathbf{C} \longrightarrow \mathbf{C}$. Let us identify $\mathbf{C}$ with $\mathbf{R}^2$ in the usual manner. Let U be a domain in $\mathbf{C}$, and let $f \colon U \longrightarrow \mathbf{C}$ be a smooth map. For every $z \in \mathbf{C}$ the tangent map $d_z f$ can be considered as a real linear map $\mathbf{R}^2 \longrightarrow \mathbf{R}^2$. Let us define the complex dilatation of f at z as the point

$$\mu_f(z) = \mu_{d_z f}$$

of the open unit disc in $\mathbf{R}^2 = \mathbf{C}$ considered as a complex number.

4.5. Theorem. *Let U be a domain in $\mathbf{C}$ and let $f \colon U \longrightarrow \mathbf{C}$ be a smooth map. If f is non-degenerate at a point $z \in U$, then*

$$\mu_f(z) = \overline{\partial} f(z) / \partial f(z).$$

In other words, the definition of the complex dilatation given in 4.4 agrees with the classical one.

Proof. It is well known that the classical complex dilatation $\mu = \overline{\partial} f(z)/\partial f(z)$ of f at z is characterized by the following two properties (cf. [A], Section I.A).

(i) The differential $d_z f$ maps infinitesimal circles around z into infinitesimal ellipses around $f(z)$ with the ratio of the major axis to the minor one equal to

$$(8) \qquad \gamma = \frac{1 + |\mu|}{1 - |\mu|},$$

where γ is the ratio of the major axis to the minor one.

(ii) The absolute value of the directional derivative $\partial_\sigma f(z)$ in the direction having the angle σ with the positive real half-line is maximal when $2\sigma = \arg\mu$.

This characterization matches the properties of $\mu_f(z)$. Indeed, (8) is equivalent to

$$|\mu| = \frac{\gamma - 1}{\gamma + 1}.$$

The directional derivative $\partial_\sigma f(z)$ is obviously maximal when the direction with the angle σ is mapped to the major axis of the image ellipse, i.e. when σ is equal to θ from the subsection 4.3. By combining these observations with (7), we see that $\mu_f(z)$ has the properties (i) and (ii). The theorem follows. ■

References

[A] L.V. Ahlfors, *Lectures on quasiconformal mappings*, Van Nostrand, Princeton, New Jersey, 1966.

[B] M. Berger, *Geometry II*, Springer-Verlag, Berlin, Heidelberg, New York, 1987.

[Ca] J.W.S. Cassels, *Rational quadratic forms*, Academic Press, London and New York, 1978.

[Cl] C.H. Clemens, *A scrapbook of complex curve theory*, Plenum Press, New York and London, 1980.

[H] J.H. Hubbard, The monodromy of projective structures, *Riemann surfaces and related topics: Proceedings of the 1978 Stony Brook Conference*, I. Kra and B. Maskit, Editors, Annals of Mathematics Studies No. 97, Princeton University Press, Princeton, New Jersey, 1981, pp. 257–275.

[T] W.P. Thurston, Zippers and univalent functions, *The Bieberbach conjecture: Proceedings of the symposium on the occasion of the proof*, A. Baerstein, D. Drasin, P. Duren, and A. Marden, Editors, Mathematical Surveys and Monographs, No. 21, AMS, 1986, pp. 185–197.

December 21, 1995

November 6, 2016 (Reformatting and minor edits)

https://nikolaivivanov.com

E-mail: nikolai.v.ivanov@icloud.com