Journal for Geometry and Graphics
Volume 23 (2019), No. 1, 65–83.

An (Isometric) Perspective on Homographies

Annalisa Crannell, Marc Frantz, Fumiko Futamura

Department of Mathematics, Franklin & Marshall College
Lancaster, PA 17604-3003, USA
email: annalisa.crannell@fandm.edu

Abstract. We revisit the classical theorem stating that every non-affine homography $\mathbf{h}\colon \mathbb{E}^2 \to \mathbb{E}^2$ defined on the extended Euclidean plane can be written as the composition of an isometry with a perspective collineation. We show there are exactly four such decompositions, such that the isometry is orientation preserving or reversing, and the perspective collineation has positive or negative cross ratio. This decomposition gives us a natural way to describe a cross-ratio of $\mathbf{h}$ itself, and also to describe a measure we call *anamorphic distance distortion* at points in $\mathbb{E}^2$; we show this distortion is invariant along circles in the image space. Applications to analyzing perspective art include location of the camera in the domain plane and interpretation of anamorphic art as a function of viewing distance.

Key Words: homography, collineation, cross-ratio, anamorphism, perspective
MSC 2010: 51N05, 51A05, 00A66.

1. Introduction

A planar *homography* (sometimes also called a *projective collineation*) is a function that takes points to points and lines to lines. Two standard examples of homographies include affine transformations and perspective collineations; these two examples are, in fact, the *only* examples, up to composition with isometry.

A classical result of projective geometry states that every non-affine homography is the composition of an isometry with a perspective collineation. It follows that there are many aspects of this decomposition that are uniquely determined: each non-affine homography comes with a unique *principal point*, a unique *base point*, a unique *principal distance*, and a unique *base distance*. Moreover, we show that the homography can be decomposed as an isometry with a perspective collineation in exactly four ways (see for example Figure 8), such that the isometry is orientation preserving or reversing, and the perspective collineation has positive or negative cross ratio. (In this sense, we can decompose the homography uniquely as, say, an orientation preserving isometry composed with a positive-cross-ratio perspective collineation).

ISSN 1433-8157/$ 2.50

We give applications of this theorem to perspective art and photographic images. For example, the principal point and base point and the associated distances correspond to the location of the artist, both in relation to the scene being portrayed, and also in relation to the canvas. The question of locating the camera for a given photographic image is a common one (see for example references [3],[5], [20], and [23]); we describe a geometric, ruler-and-compass construction in the ground plane (that is, the domain) for locating the camera. In addition, we demonstrate that, for every non-affine homography, we can compute the *anamorphic distance distortion* at various points in the plane; this measurement is closely related to standard perspective art notions of anamorphism.

Here is an overview of the paper. In Section 2, we present definitions that give us a geometric structure for analyzing homographies. In Section 3, we use the underlying structure of the homography to give a geometric construction for locating the camera.

Section 4 presents our main theorem: every same-plane non-affine homography can be decomposed as a perspective collineation composed with an isometry in exactly four ways. We use "unfolding" arguments like those of Brook TAYLOR to motivate a notion of cross-ratio for these homographies that corresponds in a natural way to the cross-ratios of the corresponding perspective collineations.

Section 5 defines the *anamorphic distance distortion* measure, which depends on the distance between the principal vanishing point and image of the point; we conclude with two examples from perspective art that further motivate the descriptiveness of this measure.

2. Principal point, line, and distance, together with base point, line, and distance

Properties of the principal and base lines

We will work in $\mathbb{E}^3$, extended Euclidean space. We choose planes $\mathbb{E}^2_b, \mathbb{E}^2_v \subset \mathbb{E}^3$ (possibly identical), with the symbols b and v chosen to evoke "base plane" and "vertical picture plane" (and for some objects in this latter plane, "vanishing points").

The following theorem can be found in various guises in multiple sources (see, for example, [2], [7], [8], [10], [16], and [21]), often stated as "every quadrangle is in perspective with a square."

Theorem 1. *Given a non-affine homography* $\mathbf{h}: \mathbb{E}^2_b \to \mathbb{E}^2_v$, *we can write* $\mathbf{h} = \mathbf{i} \circ \mathbf{p}$ *as the composition of a perspective collineation* $\mathbf{p}: \mathbb{E}^2_b \to \mathbb{E}^2_v$ *and an isometry* $\mathbf{i}: \mathbb{E}^2_v \to \mathbb{E}^2_v$.

Our goal is to apply this theorem to interpreting images, using both same-plane and two-plane interpretations of perspective collineations. We discuss the former interpretation in Section 4, but our main focus is the *standard form*, illustrated in Figure 1, in which we consider $\mathbb{E}^2_b$ as the horizontal ground plane and $\mathbb{E}^2_v$ as a vertical canvas; the line v is the *horizon*, meaning that each collection of parallel lines in $\mathbb{E}^2_b$ "converges" (or rather, intersects) at a *vanishing point* on v. Figure 1 shows a perspective collineation and not a general homography; but from Theorem 1 it follows that every non-affine homography can be interpreted as a perspective mapping of this form followed by some isometric transformation of the picture plane.

Accordingly, we impose a structure on our planes using the following important lines:

- the *base line* $b = \mathbf{h}^{-1}(i_v)$, the preimage in $\mathbb{E}^2_b$ of the ideal line $i_v \in \mathbb{E}^2_v$, and
- the *vanishing line* $v = \mathbf{h}(i_b)$, the image in $\mathbb{E}^2_v$ of the ideal line $i_b \in \mathbb{E}^2_b$.

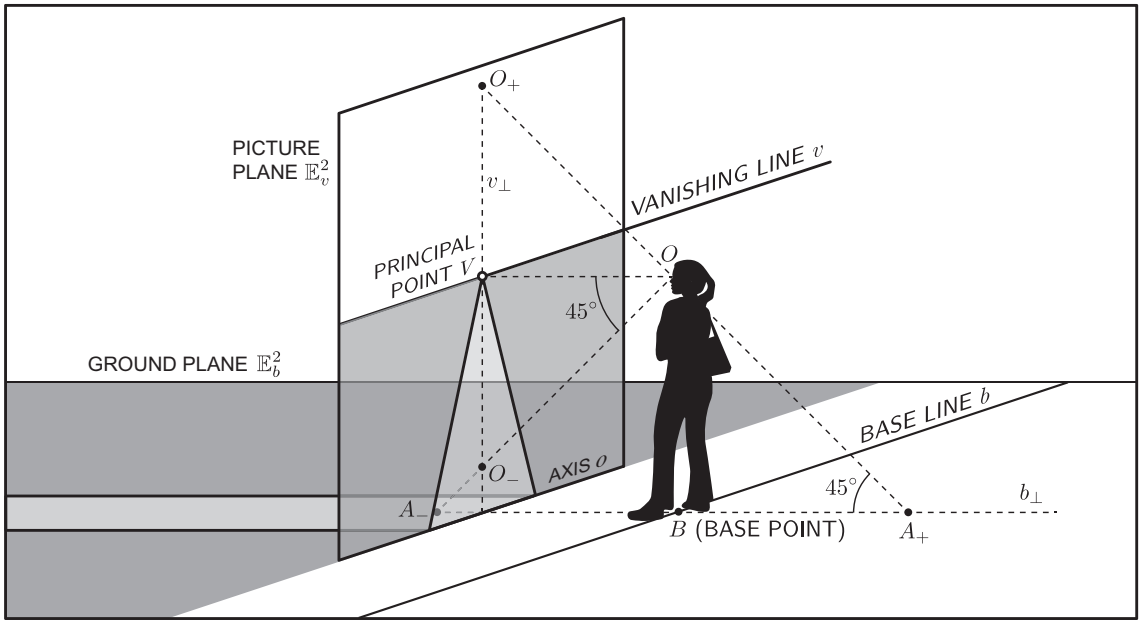

Figure 1: A perspective art interpretation of a homography mapping the ground plane to a vertical picture plane, including other points and lines introduced in this section.

We will call the lines perpendicular to the base line b *depth lines*. Their images intersect at a common point that we call $V \in v$, the *principal point* of the homography $\mathbf{h}$. We designate by $b_\perp$ the unique line perpendicular to b whose image $v_\perp = \mathbf{h}(b_\perp)$ is perpendicular to v and passing through V. The intersection $B = b \cap b_\perp$ we will call the *base point* of the homography.

The base line and vanishing line play important roles in standard correspondences between lines in $\mathbb{E}^2_b$ and their images in $\mathbb{E}^2_v$. In particular, we will often refer to the correspondences listed in the following lemma (see Figure 2).

Lemma 2. *Let $\mathbf{h} : \mathbb{E}^2_b \to \mathbb{E}^2_v$ be a non-affine homography with base line b, base point B, principal line v, and principal point V. Suppose $\ell, k \in \mathbb{E}^2_b$ are both ordinary lines. Then*

a) $\ell \parallel k \iff (\mathbf{h}(\ell) \cap \mathbf{h}(k)) \in v$;

b) $\ell \parallel b \iff \mathbf{h}(\ell) \parallel v$;

c) $\ell \perp b \iff \mathbf{h}(\ell) \ni V$; *and*

d) $\ell \ni B \iff \mathbf{h}(\ell) \perp v$.

A perspective art interpretation of statements (b) and (c) form the basis of one-point perspective drawings: statement (b) tells us that lines on the ground that are parallel to the vertical canvas have images on the canvas that are parallel to the horizon, and statement (c) says that depth lines appear to converge to the vanishing point V. Statement (d) tells us the not-entirely-intuitive fact that lines that pass directly under the artist's feet have images that are vertical line segments on the canvas (see images of the two roads in Figure 3).

Points and lines of invariance

Let us assume an orientation on each of the lines b, $b_\perp$, v and $v_\perp$, so that we can describe signed distances $|*|_{\mathbb{B}}$ and $|*|_{\mathbb{V}}$ along these lines (that is, we have $|V_\mu V|_{\mathbb{V}} = -|V V_\mu|_{\mathbb{V}}$, etc.).

Without loss of generality, we impose an orientation on the respective planes such that a counter-clockwise 90° rotation brings the positive orientation of b to that of $b_\perp$, and similarly

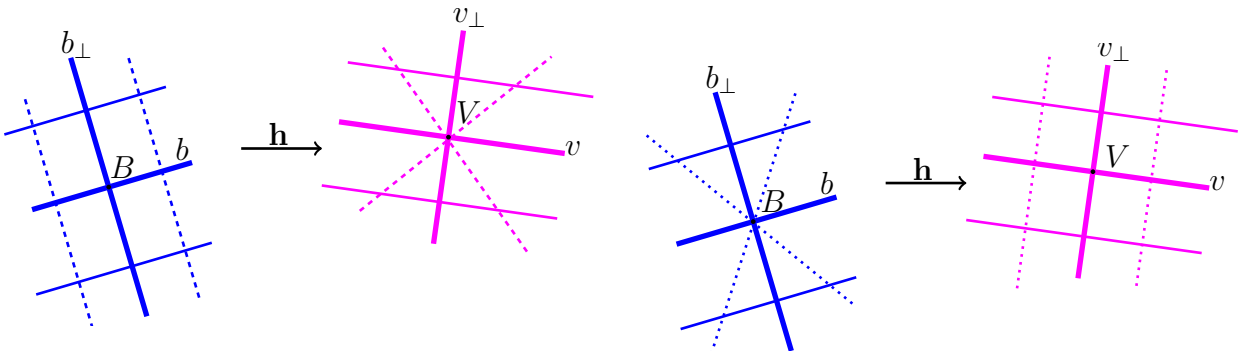

Figure 2: (left) Lines parallel to the base line b have images parallel to the principal line v; lines perpendicular to b have images that pass through the principal point V. (right) Lines that pass through the base point B have images perpendicular to the principal line v.

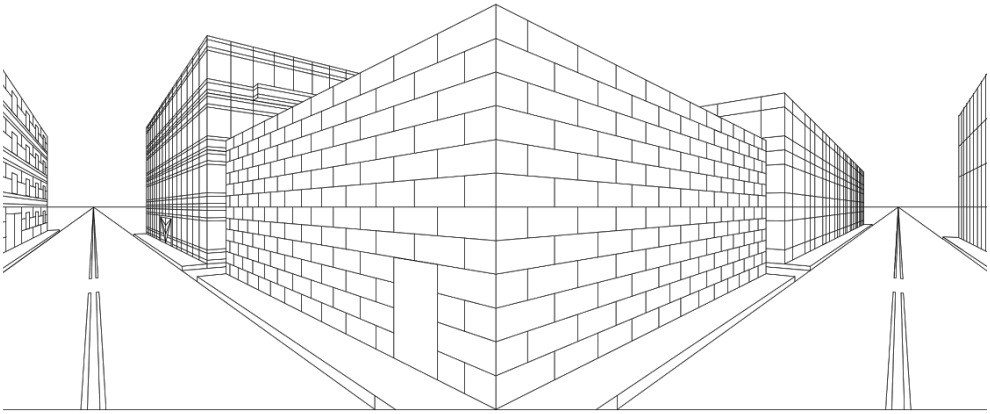

Figure 3: Gega AGULASHVILI's homage to a question that begins, "An artist stands at the intersection of two roads". The images of the center line of each road is a vertical line segment, perpendicular to the horizon, illustrating statement (d) of Lemma 2.

with v and $v_\perp$. Thus, we have introduced right-handed cartesian coordinate frames. Note that observing the base plane or picture plane from the opposite direction reverses the orientation.

A consequence of Theorem 1 tells us the homography $\mathbf{h}$ has exactly two angle-preserving points (see [8], [10], [16]). We can accordingly denote these points

$$O_+ = \mathbf{h}(A_+), \quad \text{and} \quad O_- = \mathbf{h}(A_-)$$

in such a way that $\mathbf{h}$ preserves angles and orientation at A_+; consequently $\mathbf{h}$ preserves angles and reverses orientation at A_-. The definition does not assume or require that A_+ and O_+ are on the positive sides of their respective lines; note, however, that reversing the orientation of both b and $b_\perp$ leaves the resulting slope of lines in $\mathbb{E}^2_b$ unchanged, and likewise for $\mathbb{E}^2_v$. We will therefore assume in all that follows that we have chosen the orientations of the four main lines so that the distances $\delta_b = |BA_+|_{\mathbb{B}}$ and $\delta_v = |VO_+|_{\mathbb{V}}$ are positive.

Similarly, Theorem 1 implies the homography $\mathbf{h}$ has exactly two distance-preserving lines; we denote these by

$$o_+ = \mathbf{h}(a_+) \quad \text{and} \quad o_- = \mathbf{h}(a_-).$$

in such a way that $\mathbf{h}: a_+ \to o_+$ preserves absolute distances and orientation; $\mathbf{h}: a_- \to o_-$ preserves absolute distances while reversing orientation.

Figure 1 illustrates the angle-preserving points in the particular instance in which the homography is a perspective mapping via the center O and the axis $o := o_+ = a_+$. Note that $o \parallel b \parallel v$, and that

$$\delta_b = |BO|_{\mathbb{V}} = |BA_+|_{\mathbb{B}} = |A_-B|_{\mathbb{B}} = |o_+v|_{\mathbb{V}} = |vo_-|_{\mathbb{V}} \tag{1}$$

and

$$\delta_v = |VO|_{\mathbb{B}} = |VO_+|_{\mathbb{V}} = |O_-V|_{\mathbb{V}} = |a_+b|_{\mathbb{B}} = |ba_-|_{\mathbb{B}}. \tag{2}$$

Equations 1 and 2 hold generally for homographies, as in Figure 4.

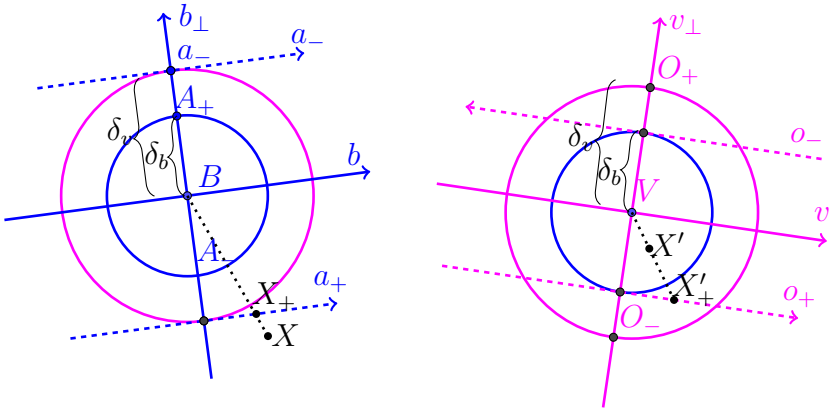

Figure 4: The existence of distance-preserving lines for $\mathbf{h}$. The points X, X_+, X', and X'_+ illustrate the proof of Theorem 6.

The proof of the following lemma is straightforward; we leave it to the reader.

Lemma 3. *Let $\mathbf{h}: \mathbb{E}^2_b \to \mathbb{E}^2_v$ be a non-affine homography with distances $\delta_b = |BA_+|_{\mathbb{B}}$ and $\delta_v = |VO_+|_{\mathbb{V}}$. Choose an ordinary line $\ell \subset \mathbb{E}^2_b$. Then*

1. $\text{slope}_{\mathbb{B}}(\ell) = \mu$ *if and only if the point* $V_\mu = v \cap \mathbf{h}(\ell)$ *satisfies* $|VV_\mu|_{\mathbb{V}} = -\delta_v/\mu$; *and*
2. $\text{slope}_{\mathbb{V}}(\mathbf{h}(\ell)) = \mu$ *if and only if the point* $B_\mu = b \cap \ell$ *satisfies* $|BB_\mu|_{\mathbb{B}} = -\delta_b/\mu$.

Suppose $\ell \parallel b$, and let the distance $|b\ell|_{\mathbb{B}}$ between these lines be given by the orientation of $b_\perp$. Then $\mathbf{h}(\ell) \parallel v$, so the distance $|v\mathbf{h}(\ell)|_{\mathbb{V}}$ is likewise well defined, and moreover,

$$|b\ell|_{\mathbb{B}} \cdot |v\mathbf{h}(\ell)|_{\mathbb{V}} = \delta_b \cdot \delta_v.$$

We conclude with a lovely geometric interpretation that describes the principal distance in terms of images of perpendicular lines. In perspective art, this allows us to "see" the viewing distance (principal distance) as a specific altitude of a semicircle bounded by two vanishing points.

Corollary 4. *Suppose ℓ and k are perpendicular ordinary lines in $\mathbb{E}^2_b$, with ideal points I_ℓ and I_k. Then δ_v is the geometric mean of the distances $|\mathbf{h}(I_\ell)V|_{\mathbb{V}}$ and $|V\mathbf{h}(I_k)|_{\mathbb{V}}$.*

Proof. See Figure 5. Suppose ℓ has slope μ; it follows that k has slope $-1/\mu$. By Lemma 3 it follows that $\mathbf{h}(I_\ell) = V_\mu$ and $\mathbf{h}(I_k) = V_{-1/\mu}$, so that

$$\sqrt{|\mathbf{h}(I_\ell)V|_{\mathbb{V}} \cdot |V\mathbf{h}(I_k)|_{\mathbb{V}}} = \sqrt{|V_\mu V|_{\mathbb{V}} \cdot |VV_{-1/\mu}|_{\mathbb{V}}} = \sqrt{\left(-\frac{\delta_v}{\mu}\right) \cdot (-\delta_v \mu)} = \delta_v. \qquad \square$$

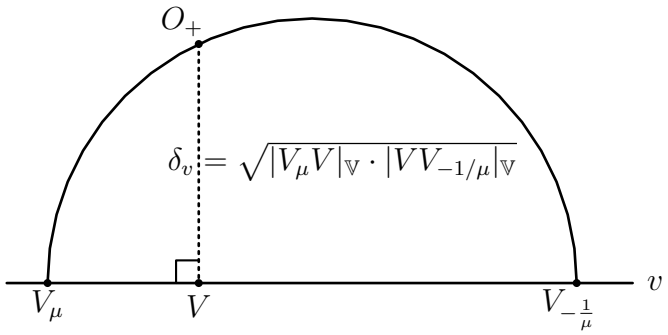

Figure 5: Illustration of Corollary 4.

3. Construction of base point, base line, and base distance

The question "where was the artist?" is a standard question in perspective art, and is usually determined by looking almost exclusively at the image (picture) plane. Here we provide a geometric construction for locating the artist directly in the base (ground) plane. The techniques avoid explicit computation, and can rely instead on straightedge-and-compass constructions, which enable us to duplicate distances and also angles.

Figure 6 shows the beginning of a geometric construction. Imagine we have a set of lines, such as train tracks (left), and their image (right). We may determine the principal point V and a principal distance $\delta_v = |VO_+|_{\mathbb{V}}$ in the standard way (see [1], [11], [12], etc).

In [14], Hartley and Silpa-Anan describe a technique for using of the angle-preserving properties of O_+ and O_- (which they call *conformal points* of the homography) to determine relative angles of lines in $\mathbb{E}_b^2$. We use their technique here to locate the base point B, as follows (see Figure 6).

Choose points $X_1, X_2 \in \mathbb{E}_b^2$ such that the line $\ell = X_1X_2$ does not pass through B (that is, the image $X_1'X_2'$ is not perpendicular to the principal line v). In $\mathbb{E}_v^2$, draw lines x_1', x_2' through X_1' and X_2' perpendicular to the principal line v; by Lemma 2 those new lines must necessarily be images of lines x_1 and x_2 that pass through B. The intersections $V_1 = x_1' \cap v$ and $V_2 = x_2' \cap v$ tell us the angles that x_1 and x_2 make with $\ell = X_1X_2$; that is, the angle $\theta = \angle V_2O_+W$ is the same as the angle between x_2 and ℓ. Therefore, we can draw lines x_1 and x_2 through points X_1 and X_2 respectively. We thereby determine $B = x_1 \cap x_2$.

Next we construct line $b_\perp$ through B; we do so easily because the angle between $b_\perp$ and ℓ is the same as the angle between $v_\perp$ and WO_+. We can accordingly construct the base line b through B, perpendicular to $b_\perp$.

From there, we can construct A_+ and A_-. To do this, we can use the Hartley/Silpa-Anan construction again: for example, we could locate the intersection $V_3 = (X_1'O_-) \cap v$, and use the resulting angle $\angle WO_+V_3$ to construct a line through X_1 that intersects $b_\perp$ at A_-.

We can interpret the distance $\delta_b = |A_-B|_{\mathbb{B}}$ as something we might call a *height factor*. That is, we imagine the artist is standing directly on top of B and places his or her eye at a height of $\delta_b = |A_-B|_{\mathbb{B}}$. We can then construct the axis o for our perspective mapping as in Figure 7. In this interpretation, the choice of placing the axis at a_+ in $\mathbb{E}_b^2$ and at o_+ in $\mathbb{E}_v^2$ implies the artist is looking toward the train tracks through a window; if instead we placed the axis at a_- and o_-, that would be like producing an image via a pin-hole camera.

We read the left side of Figure 7 as follows: an artist stands near the train tracks at point B, facing A_-, and places a canvas vertically upon the axis a_+ that is a distance δ_v from B. On the right, we see the vertical canvas, which intersects the ground line at the axis o_+ that

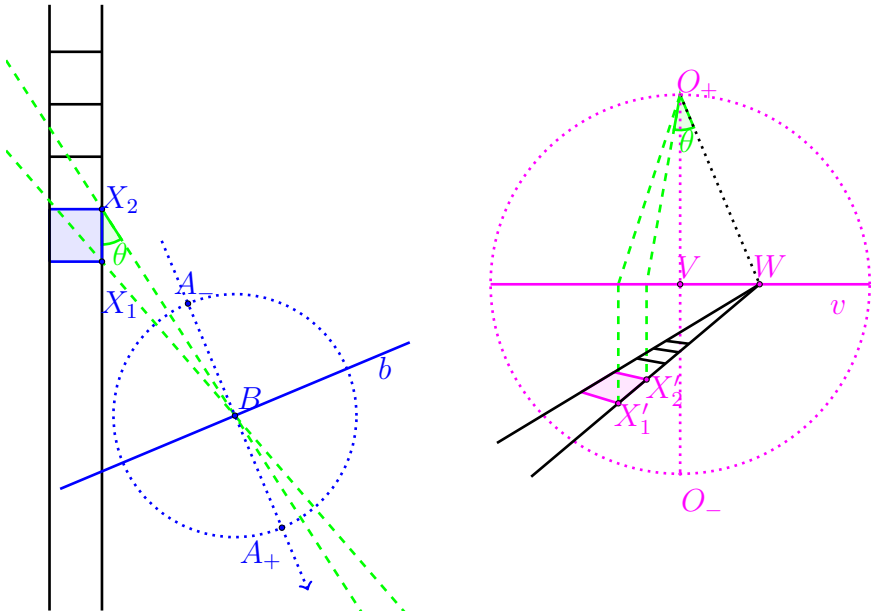

Figure 6: Constructing B and $\delta_b = |BA_+|_{\mathbb{B}}$, given an overhead view of train tracks and their image, with the principal (vanishing) line v, principal point V, and the principal distance $\delta_v = |VO_+|_{\mathbb{V}} = |O_-V|_{\mathbb{V}}$.

lies a distance δ_b from the principal point V. As above, $\delta_v = |VO_+|_{\mathbb{V}} = |ba_+|_{\mathbb{B}}$ gives us the distance from the artist to the canvas, and $\delta_b = |BA_+|_{\mathbb{B}} = |o_+v|_{\mathbb{V}}$ gives us the height of the artist's eye above the ground.

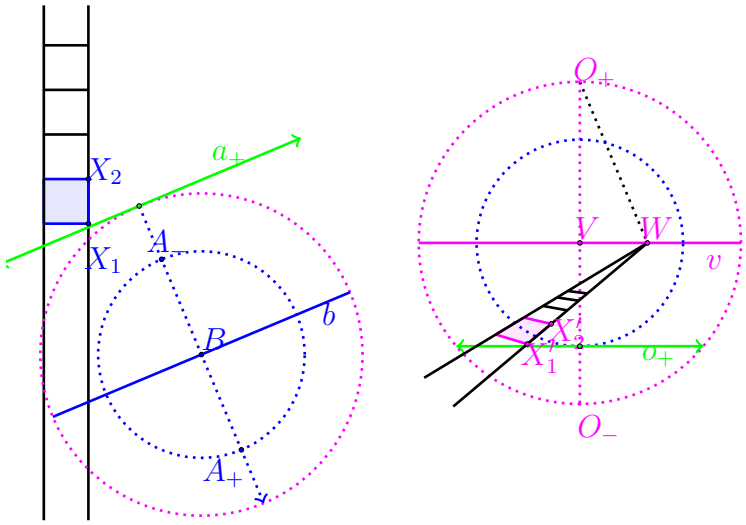

Figure 7: Constructing the axis for a perspective mapping or perspective collineation.

4. Isometries and folding

We are now ready to present our main theorem. In the remainder of the section, we will use this theorem to explore the connection between two-plane perspective maps and same-plane perspective collineations, and also to show that every homography comes with a natural

measurement that is analogous to the cross-ratio of perspective collineations.

Theorem 5. *Given a non-affine homography* $\mathbf{h}: \mathbb{E}_b^2 \to \mathbb{E}_v^2$, *there are exactly four ways we can write* $\mathbf{h}$ *as the composition of a perspective collineation* $\mathbf{p}: \mathbb{E}_v^2 \to \mathbb{E}_v^2$ *and an isometry* $\mathbf{i}: \mathbb{E}_b^2 \to \mathbb{E}_v^2$.

Figure 8 illustrates the four decompositions for one such homography. It shows four same-sized squares, each in perspective with the black/gray quadrangle $\mathbf{h}(\mathcal{S})$. The principal point V that lies on the solid black vanishing/principal line is common to all four collineations. There are two possible centers (O_+ and O_-) and two possible axes (o_+ and o_-, equidistant from the principal line v) for these collineations. Note that the blue squares are rotated 180 degrees about O_+; likewise the yellow squares are rotated 180 degrees about O_-. The orientation of the yellow squares is reversed from that of the blue squares; indeed, they are reflections though respective axes of one another.

Proof. Let us choose orientations on $\mathbb{E}_b^2$ and $\mathbb{E}_v^2$ and define points A_+, A_-, O_+, and O_- and lines a_+, a_-, o_+, and o_- accordingly.

There are four ways we choose a pair of symbols $\diamond, \star$ from the set $\{+, -\}$. By Lemma 3, for each of these four choices we can define an isometry $\mathbf{i}_{\diamond\star} : \mathbb{E}_b^2 \to \mathbb{E}_v^2$ such that $\mathbf{i}_{\diamond\star}(A_\diamond) = \mathbf{h}(A_\diamond) = O_\diamond$ and $\mathbf{i}_{\diamond\star}(X) = \mathbf{h}(X)$ for all $X \in a_\star$. Note that, for each of these isometries, we have $\mathbf{i}_{\diamond\star}(b_\perp) = v_\perp$, and so $\mathbf{i}_{\diamond\star}(I_{b_\perp}) = I_{v_\perp}$.

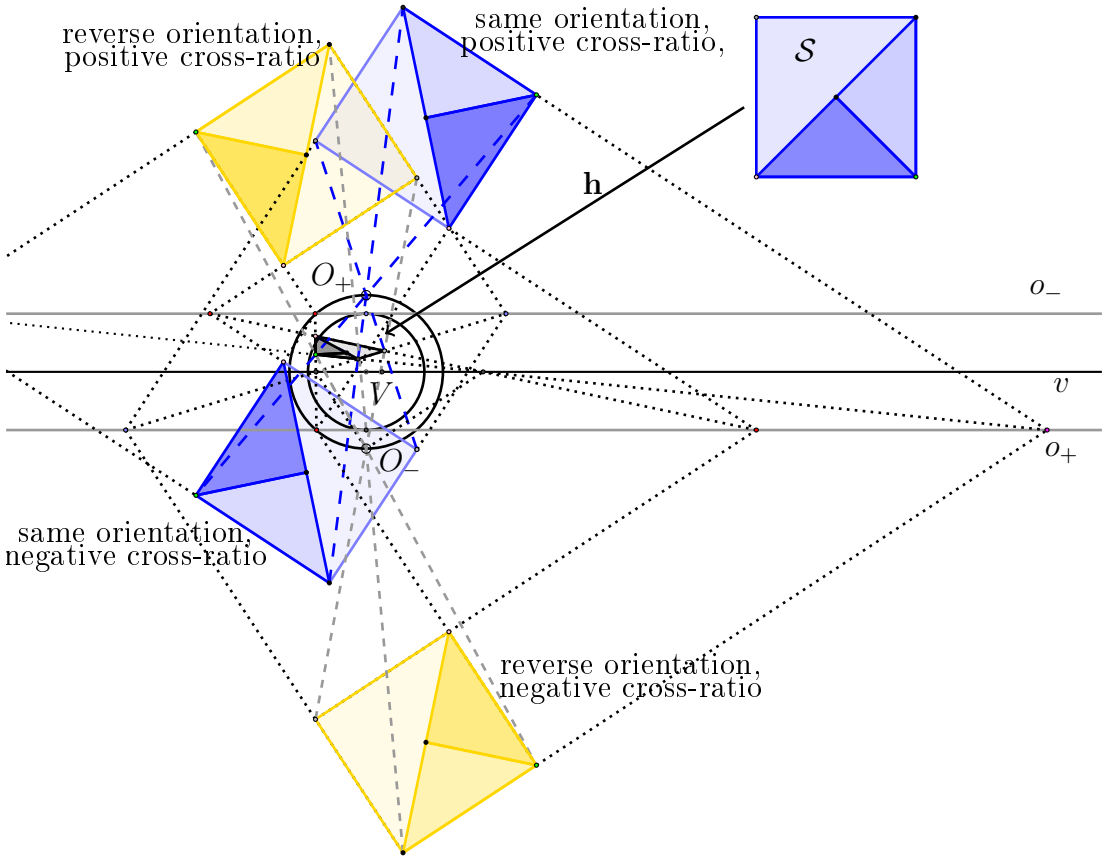

Figure 8: The square $\mathcal{S}$ is mapped by four different isometries to squares in perspective with $\mathbf{h}(\mathcal{S})$.

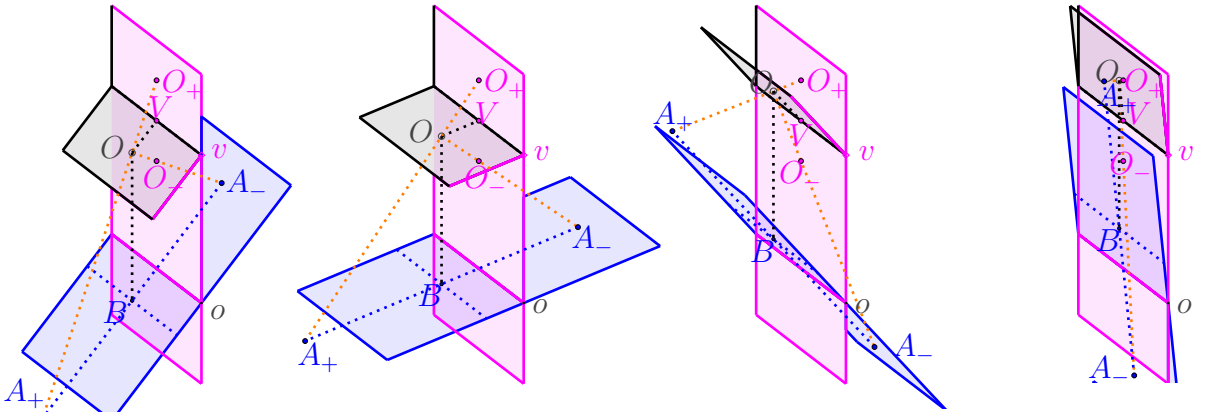

Figure 9: Folding a perspective map into two different perspective collineations. While the preimage plane is rotated about the axis o, the center O traces a circle with the axis v.

We claim that $\mathbf{p}_{\diamond\star} := \mathbf{h} \circ \mathbf{i}_{\diamond\star}^{-1}$ is a perspective collineation with center $O_\diamond$ and axis $o_\star$. Why is this so? The function $\mathbf{p}_{\diamond\star}$ fixes $O_\diamond$ because

$$\mathbf{p}_{\diamond\star}(O_\diamond) = \mathbf{h} \circ \mathbf{i}_{\diamond\star}^{-1}(O_\diamond) = \mathbf{h}(A_\diamond) = O_\diamond,$$

and similarly, $\mathbf{p}_{\diamond\star}$ fixes every point on the line $o_\star$. It follows therefore that $\mathbf{p}_{\diamond\star}(v_\perp) = v_\perp$, and that therefore

$$\mathbf{p}_{\diamond\star}(I_{v_\perp}) = \mathbf{h} \circ \mathbf{i}_{\diamond\star}^{-1}(I_{v_\perp}) = \mathbf{h}(I_{b_\perp}) = V.$$

By the Fundamental Theorem of Projective Geometry, which says that four points, no three of which are collinear, determine a homography, it follows that $\mathbf{p}_{\diamond\star}$ must be the perspective collineation with center $O_\diamond$, axis $o_\star$, and with $\mathbf{p}_{\diamond\star}(p_\infty) = v$. □

4.1. Homographies, perspective collineations, and folding

Theorem 5 allows us to describe a link between two-plane interpretations of homographies as perspective mappings and the single-plane perspective collineations. We use a "folding" analogy, described by many mathematical artists who have studied perspective mappings (see Andersen's work on Brook Taylor's perspective mathematics [1], for example). In this folding, we consider three planes, the first two of which we have seen before as in Figure 1: a ground plane $\mathbb{E}_b^2$ and a picture plane $\mathbb{E}_v^2$, which intersect along an axis o. In the description below, we will assume $o = o_+ = a_+$, but the analysis works similarly for $o = o_- = a_-$. The third plane, $\mathbb{E}_O^2$, is a plane containing the center O of the perspective mapping and the horizon (principal line) v. This implies the third plane is parallel to $\mathbb{E}_b^2$.

To connect the two-plane model to a one-plane collineation, we rotate $\mathbb{E}_b^2$ along the axis o and $\mathbb{E}_O^2$ along the principal vanishing line v into $\mathbb{E}_v^2$ so that the former planes remain parallel, as in Figure 9. If we rotate $b_\perp$ "up" onto $v_\perp$, then A_+ rotates onto O_+, and we get a collineation with center $O = A_+ = O_+$ that can be described as in Malton's diagram (Figure 10); we describe this construction in the next paragraph. If we rotate $b_\perp$ "down" (reversing orientation) onto $v_\perp$, then the perspective collineation has center $O = A_- = O_-$.

In Figure 10 drawn by Malton in his 1775 book ("A compleat treatise ... on the principles of Dr. Brook Taylor Made Clear, by Various Moveable Schemes, and Diagrams, in the Most Intelligent Manner" [18]), we see a typical perspective construction that divides the drawing

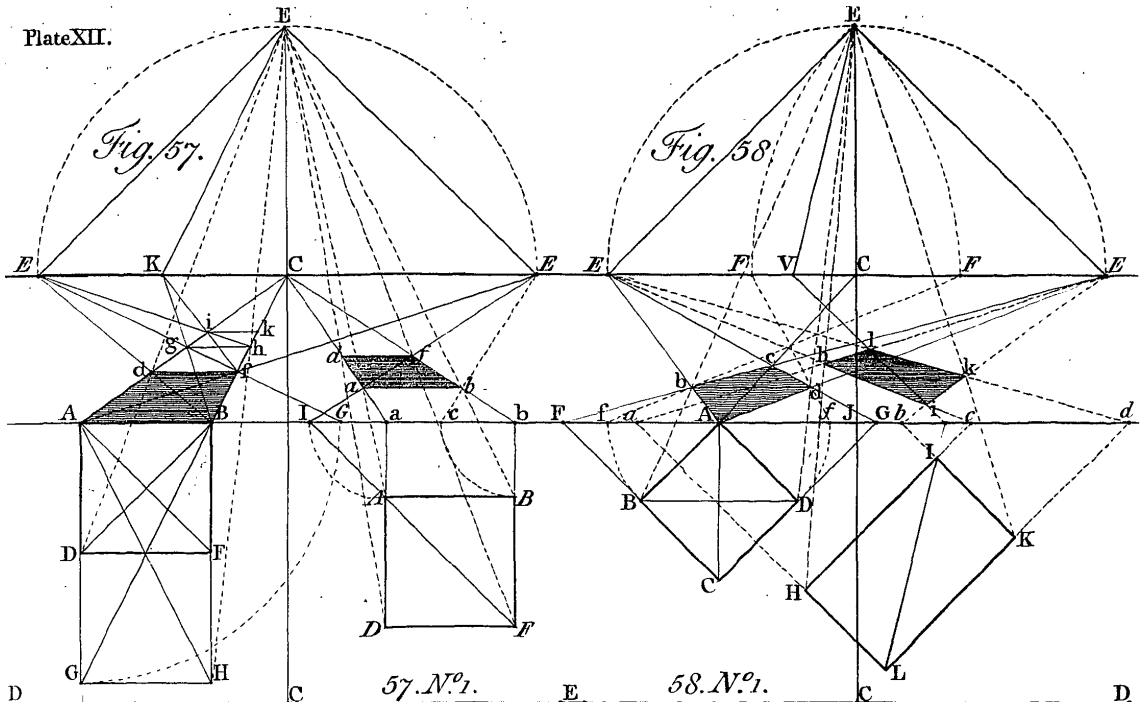

Figure 10: MALTON's diagrams of perspective constructions, after TAYLOR [18].

plane into three regions. The top region represents what we have called the plane $\mathbb{E}^2_O$ folded about the vanishing (principal) line; the point E in this figure represents what we call O, and the lines represent lines of sight to points on the vanishing line. The middle region represents images in the plane $\mathbb{E}^2_v$. The bottom region represents objects from the base plane $\mathbb{E}^2_b$. Theorem 5 notes we could have folded $\mathbb{E}^2_O$ and $\mathbb{E}^2_b$ in the other direction, but in this case all three regions would fold on top of one another, making the diagram much harder to follow. We draw a simplified version of this phenomenon in Figure 11. Note that in both foldings, the blue square and the purple quadrangle are in perspective from axis o_- and from center O_+ or O_-, and the construction methods MALTON (etc) uses are equally valid, but in the O_- case, the diagram becomes much more confusing to follow.

4.2. Cross-ratios

In this section, we describe a generalization of the cross-ratio of a perspective collineation. Suppose we have a perspective collineation $\mathbf{p}\colon \mathbb{E}^2_v \to \mathbb{E}^2_v$ with axis o and center O; then for every point $X \in P$, we can determine the points $X' = \mathbf{p}(X)$ and $X_\star = o \cap (XX')$. The cross-ratio of $\mathbf{p}$ is defined by the cross-ratio of points:

$$\langle O, X_\star; X, X' \rangle = \frac{|OX|}{|XX_\star|} \cap \frac{|X_\star X'|}{|X'O|}.$$

It is well-known (see for example [15, 629–631]) that this ratio is independent both of the choice of the point X and of the orientation $|*|$ on on the line XX'. In the definition and theorems that follow, we give an interpretation of the cross-ratio that comes from the geometric structure illustrated in Figure 4.

Definition. Given a non-affine homography $\mathbf{h}\colon \mathbb{E}^2_b \to \mathbb{E}^2_v$, we define the *oriented cross-ratio*

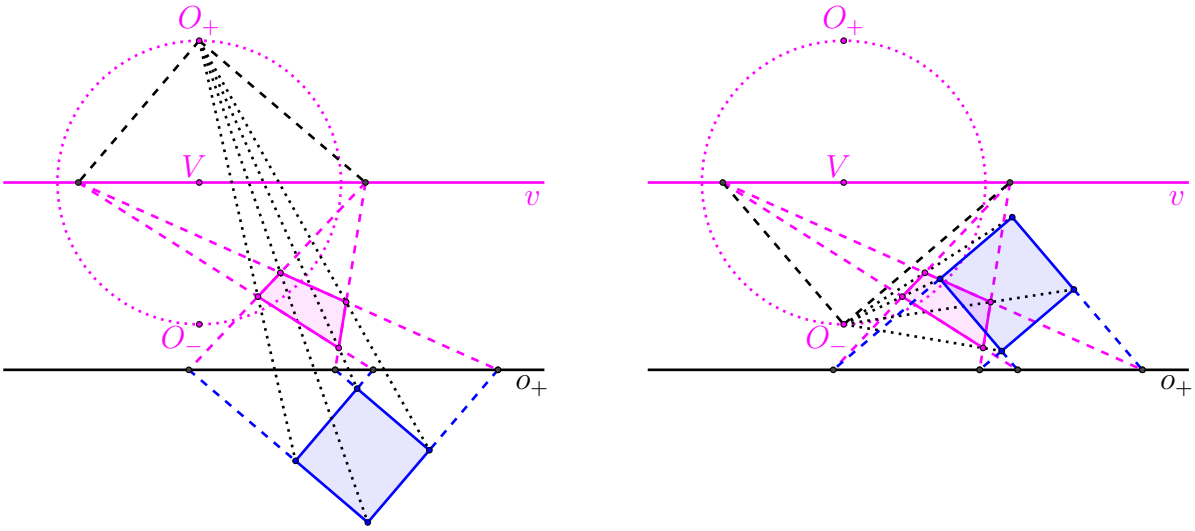

Figure 11: On the left, we fold O up to O_+; on the right, we fold O down to O_-.

of $\mathbf{h}$ to be

$$|\mathbf{h}|^\times = \frac{|BA_+|_{\mathbb{B}}}{|Ba_+|_{\mathbb{B}}} = \frac{|Vo_+|_{\mathbb{V}}}{|VO_+|_{\mathbb{V}}} = -\frac{\delta_b}{\delta_v}.$$

The next theorem describes two ways in which the definition above is a natural and meaningful generalization of the usual cross-ratio.

Theorem 6. *Given a non-affine homography $\mathbf{h}\colon \mathbb{E}^2_b \to \mathbb{E}^2_v$, choose orientations on $\mathbb{E}^2_b$ and $\mathbb{E}^2_v$ so that we may define the points A_+, O_+ (etc) and lines a_+, o_+ (etc).*

a) If we decompose the homography as $\mathbf{h} = \mathbf{p}_{\diamond\star} \circ \mathbf{i}_{\diamond\star}$ as in Theorem 5 above, then $|\mathbf{h}|^\times$ is equal to the cross-ratio of $\mathbf{p}_{\diamond\star}$ if $\mathbf{i}_{\diamond\star}$ is orientation preserving (that is, if $\diamond = \star$) and is the opposite of the cross-ratio of $\mathbf{p}_{\diamond\star}$ if $\mathbf{i}_{\diamond\star}$ is orientation reversing.

b) Choose a point $X \in \mathbb{E}^2_b$, let X_+ be the intersection of a_+ with the line XA_+, and let $X'_+ = \mathbf{h}(X_+) \in o_+$, as in Figure 4. Then

$$|\mathbf{h}|^\times = \frac{|A_+X|_{\mathbb{B}}}{|XX_+|_{\mathbb{B}}} \cap \frac{|X'_+X'|_{\mathbb{V}}}{|X'O_+|_{\mathbb{V}}}.$$

Proof. (a) We compute the cross-ratio of $\mathbf{p}_{\diamond\star}$ by choosing the point $X = I_{v_\perp}$. Note that $X' = \mathbf{p}_{\diamond\star}(I_{v_\perp}) = V$ and $X_\star = v_\perp \cap o_\star$, so

$$\langle O_\diamond, X_\star; X, X'\rangle = \frac{|O_\diamond I_{v_\perp}|}{|I_{v_\perp} X_\star|} \cap \frac{|X_\star V|}{|VO_\diamond|} = (-1) \cap \frac{|X_\star V|_{\mathbb{V}}}{|VO_\diamond|_{\mathbb{V}}} = (-1) \cap \frac{\star\delta_b}{\diamond\delta_v} = \frac{\star 1}{\diamond 1} \cap |\mathbf{h}|^\times.$$

This proves part (a).

To prove part (b), we write $\mathbf{h} = \mathbf{p} \circ \mathbf{i} = \mathbf{p}_{++} \circ \mathbf{i}_{++}$. Using the fact that $O_+ = \mathbf{i}(A_+)$ and $X'_+ = \mathbf{i}(X_+)$, we get

$$\frac{|A_+X|_{\mathbb{B}}}{|XX_+|_{\mathbb{B}}} \cap \frac{|X'_+X'|_{\mathbb{V}}}{|X'O_+|_{\mathbb{V}}} = \frac{|O_+\mathbf{i}(X)|_{\mathbb{V}}}{|\mathbf{i}(X)X'_+|_{\mathbb{V}}} \cap \frac{|X'_+X'|_{\mathbb{V}}}{|X'O_+|_{\mathbb{V}}} = \langle O_+, X'_+; \mathbf{i}(X), X'\rangle.$$

Because $X' = \mathbf{h}(X) = \mathbf{p} \circ \mathbf{i}(X)$, the expression gives us the cross-ratio of $\mathbf{p}$, which by part (a), is equal to $|\mathbf{h}|^\times$. □

Remarks. Although the cross-ratio is generally thought of as "a ratio of ratios" or a "double ratio", the previous theorem implies there is a natural interpretation of the cross-ratio of a homography as a simple ratio of two distances. In physical terms, the cross-ratio can be interpreted as the height of the camera above the ground plane divided by the distance from the camera to the vertical canvas.

5. Anamorphic distortion

In this section, we offer an argument that traditional perspective, *trompe-l'œil*, and anamorphic drawings—although they go by different names—are not fundamentally different mathematically, but rather encompass similar techniques employed in different regions of the image plane. In particular, we argue that the anamorphic distortion of a homography at a point $X \in \mathbb{E}_b^2$ is proportional (in a sense we define more explicitly below) to the distance between the image $\mathbf{h}(X)$ and the principal point V.

Hartley and Zisserman, in describing projective cameras, give kind of a measure for perspective distortion of a camera. They do this by analogy to a camera technique that is known by various names including "dolly zoom" or "Hitchcock zoom": they draw the camera away from the canvas along the principal axis, but zoom in; they then compute the resulting change in distance of the images. In our language, they let $\delta_v \to \infty$ while keeping the axis constant (that is, keeping δ_b constant). They conclude [15, p. 169],

> "From the expressions for $\tilde{\mathbf{x}}_{\text{proj}}$ and $\tilde{\mathbf{x}}_{\text{affine}}$ we can deduce
> $$\tilde{\mathbf{x}}_{\text{affine}} - \tilde{\mathbf{x}}_{\text{proj}} = \tfrac{\Delta}{d_0}(\tilde{\mathbf{x}}_{\text{proj}} - \tilde{\mathbf{x}}_0)$$
> which shows the distance between the true perspective image position $[\tilde{\mathbf{x}}_{\text{proj}}]$ and the position $[\tilde{\mathbf{x}}_{\text{affine}}]$ obtained using the affine camera approximation P_∞ will be small provided:
> (i) The depth relief (Δ) is small compared to the average depth d_0, and
> (ii) The distance of the point from the principal ray is small."

However, the distance $\tilde{\mathbf{x}}_{\text{affine}} - \tilde{\mathbf{x}}_{\text{proj}}$ ignores certain important geometric measures of distortion. We could argue, for instance, that in place of item (i) a better measure might use the *inverse* of the distance to a point: a perspective image exaggerates objects that are very close and blurs the distinctions between points that are far.

Let us look, instead, at a slightly modified version of (ii), and examine the set of points that in $\mathbb{E}_b^2$ for which the distance of *the image* of the point to the principal ray [point] is small. Indeed, points that get mapped to a circle of constant radius about V come from hyperbolas in the base plane. Lemma 7 describes these hyperbolas explicitly.

Lemma 7. *The preimage under* $\mathbf{h}$ *of* $\mathcal{C}_\rho(V)$*, a circle of radius* $\rho\delta_v$ *centered at* V*, is a hyperbola* $\mathcal{H}_\rho$ *whose foci lie on* $b_\perp$ *at distances* $\pm\frac{\delta_b}{\rho}\sqrt{1+\rho^2}$ *from* B*, and whose asymptotes pass through* B *with slopes* $\pm 1/\rho$*. The hyperbola* $\mathcal{H}_\rho$ *intersects* $b_\perp$ *at distances* $\pm\delta_b/\rho$ *from* B*.*

Proof. See Figure 12. We draw the lines tangent to the circle that are parallel to v and $v_\perp$. By Lemma 3, the lines parallel to $v_\perp$ have preimages containing B with slopes $\pm 1/\rho$; it is evident that these lines are the asymptotes of the hyperbola. By the same lemma, the lines parallel to v have preimages that are parallel to b at distances $\pm\delta_b/\rho$ from B. These lines therefore intersect the asymptotes at distances $\pm\delta_b$ from $b_\perp$.

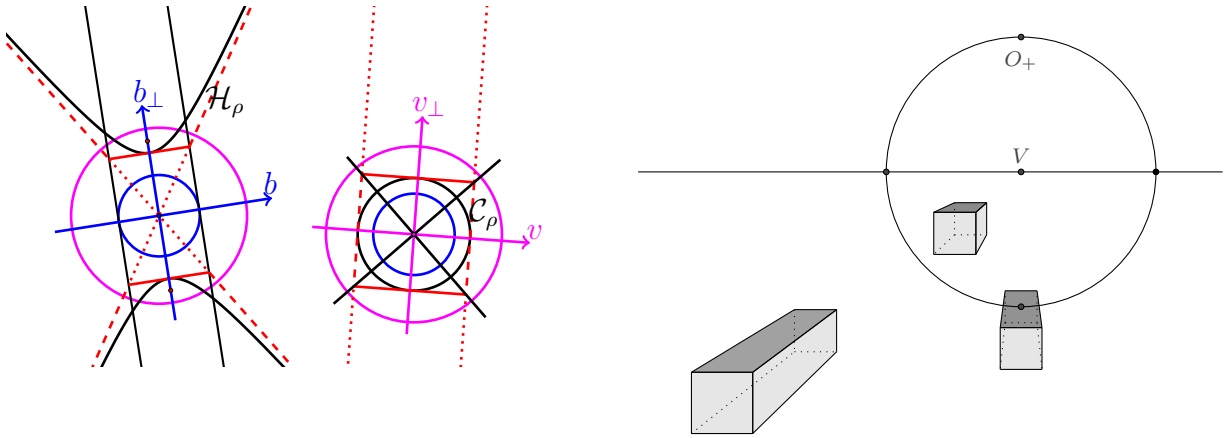

Figure 12: (left) A hyperbola $\mathcal{H}_\rho$ with axes passing through B with slopes $\pm 1/\rho$ maps under $\mathbf{h}$ to a circle of radius ρ. (right) Three images of cubes. The anamorphic distortion of the tops and sides of the cubes increases with distance from the principal point V.

It follows that the foci of $\mathcal{H}_\rho$ lie on $b_\perp$ at distances

$$\pm\sqrt{\left(\frac{\delta_b}{\rho}\right)^2 + \delta_b^2} = \pm\frac{\delta_b}{\rho}\sqrt{1+\rho^2}$$

from B. □

Now we return to the Hartley-Zisserman zoom in our setting. If we think (as earlier) of $\mathbb{E}_b^2$ as the ground plane and V as a vertical picture plane, then lines of sight from the center O to points on this hyperbola $\mathcal{H}_\rho$ all have slope ρ compared to the line OV. That is, if we move the center O to infinity, then the slope of each of these lines of sight changes by the same magnitude—from ρ to 0.

This argument gives us a two-plane way to understand that anamorphic distortion, in the sense of distortion of slope along lines of sight, is constant along circles centered at V, and that ρ itself seems to measure the distortion well. But there is a one-plane, distance distortion, way to see this distortion even more clearly.

Let us consider a collection of cubes of various sizes drawn in one-point perspective; Figure 12 (right) shows the images of three such cubes. Notice that the cube that is the furthest from the artist has an image closest to the principal point V, and that the cube that is closest to the artist, with an image furthest from V, seems to be stretched in a very "un-cube-like" manner. Similarly, it is clear that the apparent distortion of the "middle" cube seems to be greater than that of the "further" cube (with an image closer to V), even though it is closest to being at what Hartley-Zisserman might call "average depth".

In contrast, Figure 13 shows three images of squares that we claim have equal anamorphic distortion in a measurable sense. Each of these quadrangles has its perspective center lying on the circle of of radius $\rho\delta_v$. We have further normalized the picture such that the squares in $\mathbb{E}_b^2$ are different sizes, but the widths of the images in $\mathbb{E}_v^2$, measured across the perspective centers, are identical (that is, those black bars have identical lengths, say Ω).

While the images of the depth lines change angles—so that the quadrangles have very different shapes—what is surprisingly constant about these shapes is the lengths of the gray

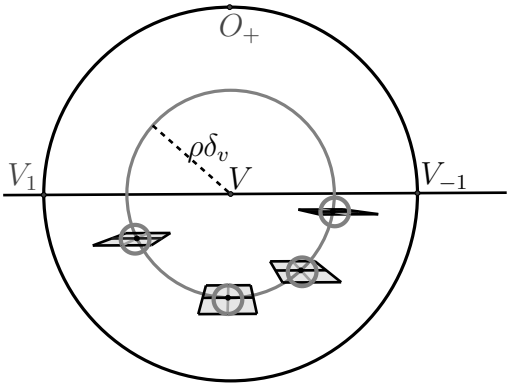

Figure 13: Images of squares whose perspective centers lie on the circle $\mathcal{C}_\rho(V)$. If the lengths of the widths are constant, then the length of the depth lines (ringed here with gray circles to make the lengths more visible) is likewise constant. See the proof of Theorem 8.

depth lines, measured again through the perspective centers (see Figure 13). The following definition and theorem make this relationship explicit.

Definition. For a non-affine homography $\mathbf{h}\colon \mathbb{E}_b^2 \to \mathbb{E}_v^2$, we define the *anamorphic distance distortion*, $\text{a.d.d.}_{\mathbf{h}}(X)$, at a point $X \in \mathbb{E}_b^2$ in the following way. Choose a square $\mathcal{S} \subset \mathbb{E}_b^2$ satisfying

- X is the center point of $\mathcal{S}$;
- the edges of $\mathcal{S}$ are parallel to b and $b_\perp$;
- the line segment through X in $\mathcal{S}$ parallel to b has an image in $\mathbb{E}_v^2$ of length Ω;
- the line segment through X in $\mathcal{S}$ parallel to b has an image in $\mathbb{E}_v^2$ of length Γ.

Then the *anamorphic distance distortion* of $\mathbf{h}$ at X is defined as $\text{a.d.d.}_{\mathbf{h}}(X) = \lim_{\Omega\to 0} \Gamma/\Omega$.

Theorem 8. *Let $\mathbf{h}\colon \mathbb{E}_b^2 \to \mathbb{E}_v^2$ be a non-affine homography as above. Suppose the point $X \in \mathbb{E}_b^2$ lies on the hyperbola $\mathcal{H}_\rho$ whose image in $\mathbb{E}_v^2$ is $\mathcal{C}_\rho(V)$, the circle of radius $\rho\delta_v$ about V. Then*

$$\Gamma = \frac{4\rho\delta_v^2\Omega}{4\delta_v^2 - \Omega^2}.$$

Therefore, $\text{a.d.d.}_{\mathbf{h}}(X) = \rho$.

Proof. We prove a somewhat stronger version of the theorem, by demonstrating that the depth-to-width ratio holds also for squares with one corner on the hyperbola. That is, we break $\mathcal{S}$ into quarters and prove a similar result for each of the four quarters.

Consider a quarter-square of $\mathcal{S}$, as in Figure 14 such that X lies on the edge closest to b (that is, $\mathbf{h}(X)$ lies on of the edge farthest from v). The width of that edge is therefore $\omega = \Omega/2$. Then we claim that the length γ of the other edge through X does not depend on the location of X along the circle.

Why is this so? See Figure 14. By similar triangles, we have

$$\frac{\omega}{\gamma} = \frac{\delta_v}{\rho\delta_v - \gamma}.$$

It follows that $\gamma = \dfrac{\rho\delta_v\omega}{\omega + \delta_v}$.

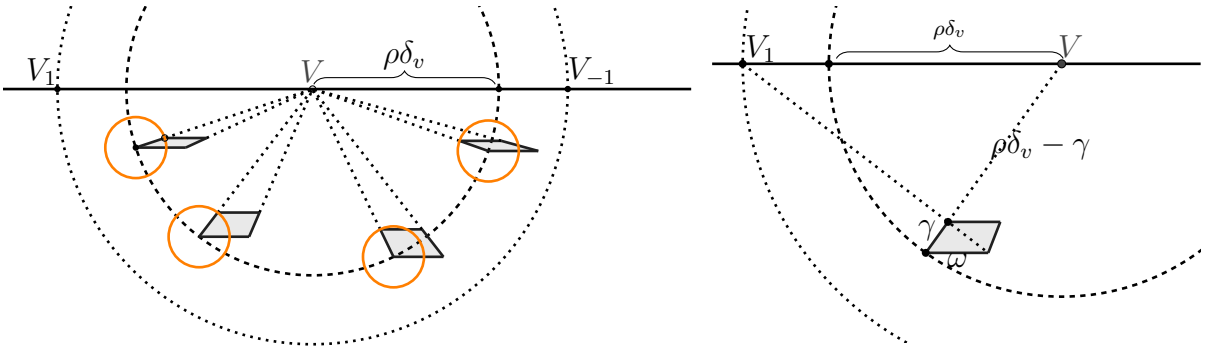

Figure 14: (left) The circles are centered on the front left corner and pass through the back left corner of the images of squares. The circles are all the same size; that is, the length of the left edge of each image is invariant with location on the circle. (right) By similar triangles, $\gamma/\omega = (\rho - \gamma)/1$, so $\gamma = \rho\omega/(\omega + 1)$.

Notice that as $\omega \to 0$ (that is, as we are drawing squares infinitesimally small), we have

$$\frac{\gamma}{\omega} = \frac{\rho\delta_v}{\omega + \delta_v} \to \rho\,;$$

that is, the anamorphic distortion of a square along the circle of radius $\rho\delta_v$ about the primary point is exactly ρ.

A similar argument yields the opposite quarter of the square has an image with one side length ω and the other side of length $\gamma' = \frac{\rho\delta_v\omega}{\delta_v - \omega}$. That gives us the length of the gray lines as

$$\Gamma = \gamma + \gamma' = \frac{2\rho\delta_v^2\omega}{\delta_v^2 - \omega^2} = \frac{4\rho\delta_v^2\Omega}{4\delta_v^2 - \Omega^2},$$

so that

$$\frac{\Gamma}{\Omega} = \frac{\gamma + \gamma'}{2\omega} = \frac{4\rho\delta_v^2}{4\delta_v^2 - \Omega^2} \to \rho$$

as $\Omega \to 0$. In other words, ρ gives us a tangible measure of the change in shape (the *anamorphosis*) at points along the circle of radius $\rho\delta_v$. □

5.1. Two examples

We conclude with two examples that give applications of Theorem 8 to traditional perspective and anamorphic art. Figure 15, our first example, depicts two well-known perspective drawings: the anamorphic 1546 painting *King Edward VI* by William Scrots (left) and the 1514 engraving *St. Jerome in His Study* by Albrecht Dürer. Although the Scrots painting seems "distorted" and Dürer etching seems "realistic", both pieces are excellent and largely correct examples of perspective from a single viewpoint. To see the head of King Edward undistorted (so that, for example, the ellipses appear to be circles), a viewer must stand to the extreme right of, and fairly close to the plane of, the painting. The primary vanishing point of Dürer's etching is also to the right, and the viewing distance is likewise known to be uncomfortably close (see for example [19, p. 181] or [6]), yet the picture still looks "good" seen from other locations. Why then does one image seem distorted and the other not?

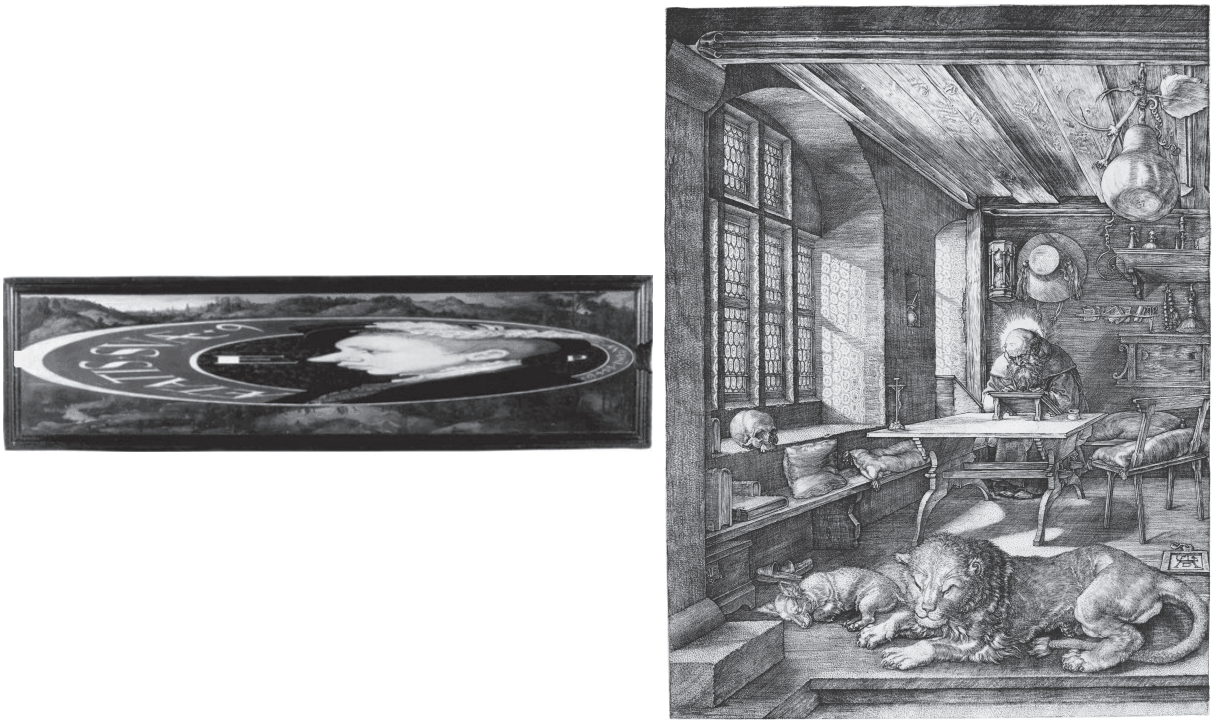

Figure 15: *King Edward VI* (from [22]) and *St. Jerome in His Study* (from [9]).

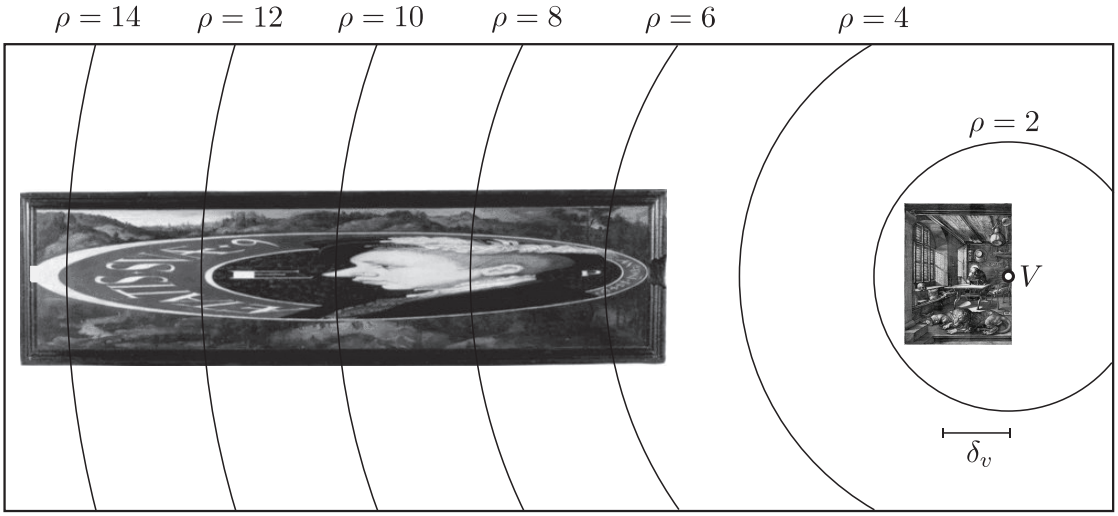

Figure 16: *King Edward VI* and *St. Jerome in His Study*, measured by ρ.

In Figure 16, we scale and translate the images to have the same primary distance δ_v and the same primary vanishing point V; observe that SCROTS's anamorphic painting lies entirely in a different region of the plane $\mathbb{E}_v^2$ than DÜRER's more traditional engraving.

We can see the phenomenon again in a second example, which compares the skull from DÜRER's *St. Jerome* with the famous anamorphic skull from Hans HOLBEIN THE YOUNGER's *The Ambassadors* (see Figure 17). These images were painted almost contemporaneously (*St. Jerome* in 1514 and *The Ambassadors* in 1533). Indeed HART and ROBSON [13] note that DÜRER "frequently depicted a skull at the base of Christ on the cross in his woodcuts" and that he was "an important influence on HOLBEIN". Their paper gives computer models for rendering the undistorted skull, using the viewing angle 9.4°. This viewing angle implies that

Figure 17: (a) The skull from DÜRER's *St. Jerome;* (b) HOLBEIN's *The Ambassadors* with a distorted skull; (c) HOLBEIN's skull, seen from the extreme right edge of the painting.

Figure 18: Measuring height-to-width ratio in the undistorted and anamorphic skulls.

the distance from the principal point V to the center of the skull is approximately $6\delta_v$ (that is, $\rho = 6.04$). In contrast, DÜRER's skull is relatively undistorted with a distance from the principal point of approximately $\rho \approx 1$.

The undistorted skull from HOLBEIN's painting is nearly equal in height to width (see Figure 18(a)). In the undistorted skull, the ratio of width to height is 90:85. In the anamorphic version, however, those lines have images with ratio 130:20. That is, the anamorphic distance distortion is approximately

$$\frac{130}{20} \cdot \frac{85}{90} \approx 6.14,$$

closely matching the HART-ROBSON estimates for ρ.

Acknowledgments

The authors thank Rolfdieter FRANK for bringing the book [2] to our attention, as well as for indicating that it is impossible to replace "similarity" with "isometry" precisely in those cases when the homography is an affine homography that takes a unit circle to an ellipse for which the semi-axes a and b satisfy $a = b$, $a \leq 1 \leq b$, or $b \leq 1 \leq a$. The authors also thank the referees for careful reading and many helpful suggestions.

References

[1] K. Andersen: *Brook Taylor's Work on Linear Perspective*. Springer-Verlag, New York 1992.

[2] H. Brauner: *Lehrbuch der Konstruktiven Geometrie*. Springer, Vienna, New York 1986.

[3] K. Byers, J. Henle: *Where the Camera Was*. Math. Mag. **77**/4, 251–259 (2004).

[4] H.S.M. Coxeter: *Projective Geometry*. 2nd ed., Springer, New York 2003.

[5] A. Crannell: *Where the Camera Was, Take Two*. Math. Mag. **79**/4, 306–308 (2006).

[6] A. Crannell, M. Frantz, F. Futamura: *Durer: disguise, distance, disagreements, and diagonals!* Math Horizons 8–11 (November 2014).

[7] A. Crannell, M. Frantz, F. Futamura: *Factoring a homography to analyze projective distortion*. Journal of Mathematical Imaging and Vision, (on-line March 30, 2019), DOI:10.1007/s10851-019-00881-4.

[8] H. Dörrie: *100 Great Problems of Elementary Mathematics*. Dover, New York 1965.

[9] A. Dürer: *St. Jerome in His Study*. 1514, from Wikimedia Commons, https://commons.wikimedia.org/wiki/File:Albrecht_D\%C3\%BCrer_-_St_Jerome_in_his_Study_-_WGA07318.jpg.

[10] W. Fiedler: *Die Darstellende Geometrie in Organischer Verbindung mit der Geometrie der Lage*. Druck und Verlag von B.G. Teubner, Leipzig 1883.

[11] M. Frantz, A. Crannell: *Viewpoints: Mathematical Perspective and Fractal Geometry in Art*. Princeton Univ. Press, Princeton/NJ 2011.

[12] R. Greene: *Determining the preferred viewpoint in linear perspective*. Leonardo **16**(2), 97–102 (1983).

[13] V. Hart, J. Robson: Hans Holbein's *The Ambassadors* (1533): A computer view of Renaissance perspective illusion. Computers and the History of Art **8**(2), 1–13 (1999).

[14] R. Hartley, C. Silpa-Anan: *Visual navigation in a plane using the conformal point*. In: R. Jarvis, A. Zelinsky (eds), *Robotics Research*. Springer Tracts in Advanced Robotics, vol. 6, Springer, Berlin, Heidelberg 2003.

[15] R. Hartley, A. Zisserman: *Multiple View Geometry in Computer Vision*. 2nd ed., Cambridge Univ. Press, New York 2003.

[16] F. Hohenberg, J.P. Tschupik: *Die geometrischen Grundlagen der Photogrammetrie*. In Jordan, Eggert, Kneissl (eds): *Handbuch der Vermessungskunde*, Bd. IIIa/1, J.B. Metzlersche Verlagsbuchhandlung, Stuttgart 1972.

[17] H. Holbein the Younger: *The Ambassadors*. from Wikimedia Commons, https://commons.wikimedia.org/wiki/File:Hans_Holbein_the_Younger_-_The_Ambassadors_-_Google_Art_Project.jpg.

[18] T. Malton: *A compleat treatise on perspective, in theory and practice; on the principles of Dr. Brook Taylor. Made Clear, by Various Moveable Schemes, and Diagrams, in the Most Intelligent Manner*. 1775, from Wikimedia Commons, https://commons.wikimedia.org/wiki/File:A_compleat_treatise_on_perspective,_in_theory_and_practice;_on_the_principles_of_Dr._Brook_Taylor_Fleuron_T100666-28.png.

[19] P. Maynard: *Drawing Distinctions*. Cornell Univ. Press, Ithaca 2005.

[20] A.C. Robin: *Photomeasurement*. Math. Gaz. **62**, 77–85 (1978).

[21] K. Rohn, W. Papperitz: *Lehrbuch der Darstellenden Geometrie*. Verlag von Veit & Comp, Leipzig 1906.

[22] W. Scrots: *Anamorphic portrait of Edward VI*. from Wikimedia Commons, `https://commons.wikimedia.org/wiki/File:Anamorphic_portrait_of_Edward_VI_by_William_Scrots.jpg`.

[23] C.E. Tripp: *Where is the camera? The use of a theorem in projective geometry to find from a photograph the location of a camera*. Math. Gaz. **71**, 8–14 (1987).

Received July 23, 2018; final form April 30, 2019