Convergence Function in Binocular Visual Space. I. A Note on Theory

THORNE SHIPLEY*

American Optical Company, Research Center, Southbridge, Massachusetts

(Received April 11, 1957)

A review and elucidation of the theory of the convergence function is presented here. In this connection, certain phenomenal aspects of binocular visual experience are noted; it is felt that any general theory must take cognizance of them. In particular, the Luneburg theory of the alley experiments is discussed in some detail, and an attempt is made to highlight the underlying and possibly equivocal assumptions. In emphasizing the distinction between pure mathematical postulates and pure phenomenal experience, we have hoped both to clarify those points at which arbitrary decisions may legitimately enter into the theory (e.g., the normalization), and to set the stage for deliberate experimentation.

A. INTRODUCTION

SOME six years ago, Rudolf Luneburg developed a new mathematical description of certain aspects of binocular visual experience.[1] The critical experimental data on which he based his formulation are the findings of Walter Blumenfeld, in 1913, that apparent parallel alleys are physically narrower than alleys which seem to be of the same width (Fig. 1).[2] Alleys which are visually parallel seem to be of uneven width and are convex inwards; equidistant-alleys are, in fact, distally converging, and also concave outwards.[3] In reformulating this fact, within the compass of an intrinsic geometry, Luneburg devised a three-dimensional bipolar coordinate system in which the angle of convergence plays an essential role. We shall call this angle γ [Fig. 2(B)]. Within the Luneburg geometry, the functional relationship between the bipolar angle γ and the polar visual distance r specifies in its entirety the transformation equation from visual space to physical space [Fig. 2(A) and Fig. 2(B)].[4] In the empirical determination of this function lies a possible three-dimensional solution of the classical Fechner problem, and a reaffirmation of the complexity of the psychophysiological isomorphism which we must invoke in accordance with Gestalt theory. It has long been held that the prototypical physiological function is logarithmic, not linear (e.g., Pieron).[5] It may be that it is non-Euclidean as well.

B. THE TRANSFORMATION OF COORDINATES

The Luneburg derivation[1a] of the general Fechnerian transformation [e.g., reference 1c, Eq. (6.1)], from the intrinsic polar visual coordinates [Fig. 2(A)] to the extrinsic bipolar physical coordinates [Fig. 2(B)], is quite different from a derivation which we might elaborate in analogy with Fechner's classical reasoning; but its general meaning is exactly the same. Given the values for $\varphi_1 - \varphi_2$, $\vartheta_1 - \vartheta_2$, and $r_1 - r_2$, we attempt to develop an equation which would enable us to predict
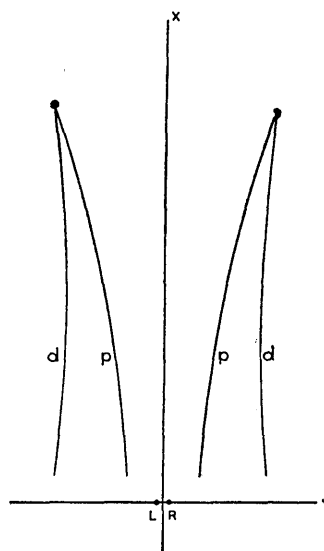
[1] R. Luneburg, (a) *Mathematical Analysis of Binocular Vision* (Princeton University Press, Princeton, 1947); (b) "Methods in binocular visual perception," in *Courant Anniversary Volume* (New York University Press, New York, 1948), pp. 215–240; (c) J. Opt. Soc. Am. **40**, 627–642 (1950).

[2] W. Blumenfeld, Z. Psychol. **65**, 241–404 (1913).

[3] The first (and only?) causal explanation of the shape of the alley curves was given by Fr. Mayer-Hillebrand [Z. Sinnesphysiol. **61**, 267–324 (1930/1931)], based on the Aubert-Foerster phenomenon [see H. Aubert and R. Foerster, Arch. Opthalmol. Graefe's **3**, 1 ff. (1857)]. However, she assumed that the alley discrepancy reported by Blumenfeld was an artifact, due to the inadvertent fixation of the end pair of lights. Blumenfeld did not accept her explanation [Z. Sinnesphysiol. **62**, 132–136 (1932)], and considered that the problem of the shape of these curves still remained.

[4] Note that this collapse of the physically discreet eyes into one single point, namely the origin of visual space—the ego-center—introduces a lacuna into our space. The distance between the two eyes, and the two eyes themselves, are not in visual space. This precludes the study of entoptic space, for example, but is not at all an unreasonable restriction from the phenomenological point of view.

FIG. 1. A schematic representation of the findings of Walter Blumenfeld (1913) that the " . . . distance-alleys [*d*] are wider, nearer to the observer, than are the parallel alleys [*p*]." *L* and *R* represent the left and right eyes, respectively.

[5] H. Pieron, *The Sensations* (Yale University Press, New Haven, 1952).

A) Polar Coordinates Visual Space

(A)

B) Bipolar Coordinates Physical Space

(B)

FIG. 2. (A) Luneburg's choice of polar coordinates in binocular visual space. O represents the polar origin, the ego-center. (B) Luneburg's choice of bipolar coordinates in physical space. L and R represent the bipolar origin, the left and right eyes, respectively.

the consequent distance and position in physical space— and vice versa.

The transformations between the two sets of coordinates (Fig. 2) may be taken as follows[6]:

Physical point Q	Visual point P
Bi-polar ϕ, θ, γ	Polar φ, ϑ, $r(\gamma)$
Cartesian x, y, z	Cartesian ξ, η, ζ

where $r^2 = \xi^2 + \eta^2 + \zeta^2$ is the radial distance in a conformal Euclidean map of visual space [as shown in Fig. 2(A)].

This chart shows that when we are given any distance, $Q_1 - Q_2$, in physical space, a corresponding distance, $P_1 - P_2$, in visual space, is uniquely specified when $r(\gamma)$ is known. Our present concern is with this latter relationship.

[6] The relations $\phi \approx \varphi$ and $\theta \approx \vartheta$ are reasonable on their face; analytical support may be found in N. Balazs and G. L. Walker. (Ms. in preparation for J. Opt. Soc. Am. Much of the present discussion will be found more fully elaborated in this paper.)

C. THE $r(\gamma)$ FUNCTION

Certain important properties of the phenomenology of distance perception have been known for many years; they are readily available to phenomenological analysis.

(a) There is nothing in our visual space to correspond with the purely intellectual concept of infinity. All visual experience is finite. Indeed, all experience is finite regardless of the sense modality. The very act of perceiving entices the stimulus into the finite realm; the two events are temporally coextensive. As long as a point is seen at all, as long as it can be said to *exist* in binocular visual space, it is seen at a finite distance, regardless of its physical distance.[7]

(b) Moreover, visual experience is bounded; all physical distances beyond a certain magnitude appear to be the same. This is true for both monocular and binocular vision, though the two bounds are by no means identical.[8]

(c) Visual space possesses a marked singularity. In an accurate phenomenology, the origin of visual space cannot be placed anywhere at will. It is the self which sees, and it is the self which must, perforce, contain the origin. Zero radial distances must always describe the strange case that the point of regard is, actually, within the head.[9]

(d) The apparent distance of an object (up to the boundary) increases regularly with the decrease in the convergence of the two eyes, though at a somewhat slower pace.[10]

---

[7] If an infinite visual experience did exist for us, then the choice of the unit would be entirely restricted; it would be the absolute radius of the pseudosphere, $r_\infty = 2/(-K)^{\frac{1}{2}}$. This has been mentioned time and again by previous workers in this area. However, such an experience does not exist and cannot exist. We suggest that all references to this possibility cease; they are entirely fanciful.

[8] A clear distinction between the various bounds of visual experience has not yet been made. [But see A. Linksz, Trans. Am. Ophthalmol. Soc. 52, 877–946 (1955).] The bound of binocular visual space is defined, physiologically, by the cessation of the significance of the angle of convergence (or of convergence innervation). This is, typically, much nearer than the distance at which, mathematically, $\gamma = 0$. Moreover, it is also, typically, much nearer than the horizon distance, itself primarily a monocular perspective experience. (This point is of relevance in explanations, for example, of the Moon illusion.) The true binocular-bound would be that distance beyond which an actually receding pin-point of light ceased to appear to recede. The reason this distance has not been determined in our laboratory is that it lies well beyond the dimensions of our darkroom.

[9] In his discussion of the phenomenology of vision, E. Hering, *Beitrage zur Physiologie* (W. Engelmann, Leipzig, 1861–1864) placed this origin at the point of momentary fixation, the so-called nuclear point; he and his students [e.g., F. P. Fischer and J. W. Wagenaar, Documental Ophthalmol. 7–8, 359–391 (1954)] have therefore to struggle with the very difficult problem of the transformation thence to the Cyclopean (ego?) center.

[10] This point, it must be emphasized, refers to a purely binocular visual environment; no monocular cues whatsoever can be permitted to enter. The apparently relevant discussion of these issues by H. E. Gruber [Am. J. Psychol. 69, 469–476 (1956)] in connection with the experimental work of A. Gilinsky [Psychol. Rev. 58, 460–482 (1951)] is so greatly interwoven with confusions between monocular and binocular environments, that it is difficult to evaluate. As a single example: "over-constancy" has never been demonstrated in an exclusively binocular environment. All the
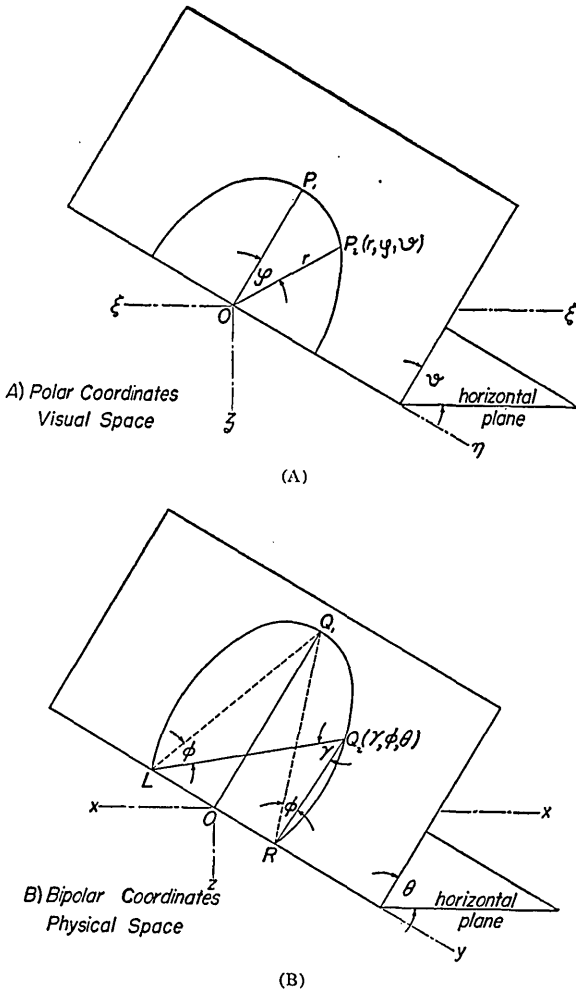
(e) There is no inherent component of visual experience which would automatically predispose us to one or another visual unit, say visual miles, feet, or inches. Therefore, in any general mathematization, whether binocular or not, the choice of the unit must remain entirely unrestricted.[7]

The more specific analytical consequences of these points have been discussed elsewhere.[6] Balazs and Walker have shown rigorously that beyond these, the $r(\gamma)$ function is most probably arbitrary, depending upon empirical and practical considerations only.

There have been two previous attempts to reduce this arbitrariness, Luneburg[1] and Hardy et al.[11] Luneburg, on the basis of some preliminary empirical evidence, considered it to be a logarithmic function $\delta \ln r/\delta\gamma = -\sigma$ such that

$$r(\gamma) = e^{-\sigma(\gamma+\mu)} = Ce^{-\sigma\gamma} \qquad (1)$$

where $C = e^{-\sigma\mu}$. In doing this, the number of parameters is limited to two, C (or μ) and σ. This is further delimited by specifying C (hence μ) as constant such that $C = 2$ (the value of r when $\gamma = 0$), the r distance to physical infinity for all observers under all visual environments. Hardy, Rand, and Rittler,[12] and Zajaczkowska[13] have made use of this restriction for their determinations of K. (K is an index of the curvature of visual space—see reference 1.) *Note that this defines the unit for all our experimental magnitudes.*[14] The determination of σ remains the only task, before the visual space of an individual may be completely described.

Not entirely satisfied with the results from the Blumenfeld alley experiments, Luneburg tried to derive the same $r(\gamma)$ function from independent evidence based upon stereoscopic fusion, of unequal left and right eye images (reference 1a, p. 51 ff). But in this he forgot one critical point: *the theory must assume equal images in both eyes.* Luneburg has often confused this issue by calling γ sometimes "bipolar parallax" and sometimes "bipolar disparity." It is never disparity. The geometry is

exclusively a geometry of nondisparate images,[15] the single two-eyed image. It is implicit that visual space is elaborated by a succession of immediate neighborhoods to the point of momentary fixation, the nuclear point. And, certainly, those peripheral images which are not embedded within the Luneburg torus of constant convergence, do not appear in the analysis. The geometry describes the distance function of the angle γ. It does not describe the depth function of binocular disparity. In all our discussions, we find it very helpful to distinguish carefully between *visual depth* and *visual distance*. The cortical fusion ("physiological congruence") of unequal ("geometrical noncongruence") retinal images for the production of visual depth certainly implies a unique non-Euclidean geometry of its own, as long noted by Woodworth.[16] But the present geometry is predicated upon the proper fusion of "identical retinal images," because we may talk only of "images" within single points [the points, R and L, of Fig. 2(B)]. It is a generalized geometry of a very specialized space indeed! Dimorphic phenomena, such as true depth perception (orthostereoscopy), aniseikonic depth distortion (parastereoscopy), and anomalous retinal correspondence, all presuppose nonconformal geometry and cannot yet be described.[17]

Another, though minor difficulty exists. The $r(\gamma)$ function proposed by Luneburg,[1c] and used by Hardy, Rand, and Rittler[12] and Zajaczkowska[13] for the determination of the Gaussian curvature K violates our third condition of singularity at the origin, where the r distance must be zero. This necessitates a slight change in the curvature values reported by them. A correct expression of the same general form would be

$$r(\gamma) = C(e^{-\sigma\gamma} - e^{-\sigma\pi}) \qquad (2)$$

where

$$r(0) = 2 = C(1 - e^{-\sigma\pi}).$$

In the other previous attempt to reduce the arbitrariness of the convergence function, Hardy et al.[11] and Blank[18] have suggested that we let $r(\gamma) = r'(\gamma - \gamma_n)$, where $r(\gamma_n)$ is the greatest visible distance, such that the "iseikonic" coordinate $\Gamma = \gamma - \gamma_n$ be used in preference to the coordinate γ. We wish to emphasize that this substitution has only apparent relevance to the arbitrariness of the convergence function; in truth, it does not affect it at all.[6]

On this basis alone, the role of their suggestion is somewhat unclear; though, in another direction, it does help to highlight a hitherto ambiguous point. Luneburg

studies of this phenomenon have included, in some uncontrolled manner, contributions of monocular, binocular, and experiential origin. To show "over-constancy" in the distance alleys would be more to the point.

[11] Hardy, Rand, Rittler, Blank, and Boeder, *The Geometry of Binocular Space Perception*, ONR Terminal Report, N6onr27119; NR143-638 (J. Schiller, Elizabeth, New Jersey, 1953).

[12] Hardy, Rand, and Rittler, A.M.A. Arch. Ophthalmol. 45, 53–63 (1951).

[13] A. Zajaczkowska, (a) Quart. J. Exptl. Psychol. 8, 66–78 (1956); (b) J. Opt. Soc. Am. 46, 514–527 (1956).

[14] No special meaning is to be ascribed to the fortuitous fact, in Eq. (1), that $r(\gamma) = r(0)r(\gamma)$. This merely means that the unit, $r(0)$, and the factor, C, are identical. This can only be true of the logarithmic function. In all other cases, where $r(\gamma) \neq r(0)r(\gamma)$, to specify $r(0)$ still specifies a unit and uniquely determines the factor, C. The specification of any $r(\gamma_n)$ value would also specify the factor, C, and the unit, provided only that $0 < r(\gamma_n) < \infty$, for this entails only that $r(\gamma) = [r(\gamma_n)r(\gamma)]/r(\gamma_n)$. The choice, by Luneburg, of $r(0)$ as a unit was made because physical infinity is a rather singular value in visual space, the ultimate binocular bound. Any other $r(\gamma_n)$ would have done as well (reference 21). The choice is exactly analogous to using miles *vs* inches.

[15] Indeed, since the theory reduces the two eyes to two points in physical space, L and R, it is not yet really proper to talk of retinal images *at all* within its compass.

[16] R. Woodworth, *Experimental Psychology* (Henry Holt, New York, 1938), p. 661.

[17] Work is in progress in this laboratory extending the theory in these directions; see also: A. Linksz, *Physiology of the Eye* (Grune & Stratton, New York, 1952), Vol. II. Independently of this, Zajaczkowska (reference 13a, p. 77) has already discussed some experimental tests of such extensions.

[18] A. A. Blank, J. Opt. Soc. Am. 43, 717–727 (1953).

has implied that the visual distance to the binocular horizon is constant for a given observer (which it most probably is), and is somehow permanently carried in his visual faculty (which it most probably is not); here it serves as a yardstick (hence a unit) against which all other distance judgments are made, *even when there are no momentarily visible points actually at the horizon.* The present suggestion to the contrary does serve to reaffirm the fact that it is the distance to the farthest point in the momentary configuration which plays this role. Visual distances are phenomenally matched off against the distance of the farthest visible point; this could, of course, actually be a point at the binocular horizon distance, but in the typical case it is nearer. What this means, in other words, is that the actual phenomenal yardstick which we are wont to use for estimating distances changes with the distance of the far point; in fact, it is exactly that distance itself (cf. Gilinsky, reference 10). It is as if we visually measure short distances in inches, and long distances in feet. The measures, of course, must all be equivalent in the end, but we do seem to have some interest in the size (manageability?) of the measuring stick.

Thus, by implication, this suggests that we should employ a normalization based on $\gamma > 0$ in place of the Luneburg normalization based on the binocular horizon $\gamma = 0$. More important, however, this encourages the complete mathematization of laboratory studies. Here we can simply obtain an estimate of the distance to the far-point (hence a geodesic magnitude) and refer, by ratios, all our other visual magnitudes to this. In the case of the horizon distance, though we have a perfect right to normalize according to it since it is also a singular distance, we can never turn our measurements into actual subjective visual units because we do not have (nor can we directly get) the observer's estimate of this distance. Only on this basis, however, can we maintain that normalization according to the horizon distance is to be discouraged. In either case, we can thoroughly explore the convergence function, independently of the final transformation to subjective units; and, in either case, the convergence function is still completely unspecified.[19] We still must learn, from the empirical evidence, what is the general form and what are the least number of parameters involved for a satisfactory fit. This latter problem is particularly important when the issue of individual prediction arises, for it would be very convenient if we could employ only one, or at most two, parameters.

The present evidence is such as to throw doubt on the unique significance of Luneburg's parameters, though the more general meaning remains. The index of curvature K is a measure of the deviation of the space from Euclidean space. Used only in this most important sense

(not in its relation to the convergence function), K will evince an *inverse* relationship to visual skills, that is eye-hand (e.g., hyperbolic-Euclidean) coordination. But, in another sense, that of Luneburg, K is also a measure of the distance of the boundary of visual space, it is the visual limit of the convergence function,[20] and in this manner it may, conceivably, *directly* reflect certain visual skills which we might call "boundary" skills: estimates of very great distances, estimates of rates of approach from afar and of rates of recession to and about the horizon, predictions of the best path to a distant goal, and so on, all judgments referring to the experience of immense distance. Thus, for small distances, the absolute magnitude of K would, necessarily, be inversely related to visual skills, while for great distances it would, presumably, be directly related. Hence, the psychophysical significance of K is highly ambiguous. At present, little more than conjecture relates K to visual skills. The other Luneburg parameter σ is an analytically arbitrary measure of distance perception. We now know that the decision to express the $r(\gamma)$ function in terms of σ, can be based on empirical considerations only.[6] And the evidence[21] suggests that equally good mathematical approximations of the data may be obtained with completely different convergence functions using basically different parameters.[11]

The present study employs the classical Hillebrand alley experiment[22] in an attempt to clarify these various issues.

D. THE GEODESIC DISTANCE

The geodesic distance D is simply a more formal mathematical term for the phenomenal distance; it is the visual distance as we actually experience it, as we estimate it in subjective visual units. We thus give a proper answer to the question: "How far away is that point of light?"

Remember, that this must be a purely binocular

---

[19] Indeed, other than pointing up the difficulties of an $r(0)$ measurement, the virtue of this substitution has not been made clear. Moreover, the further suggestion (in 11) that $r(\gamma) = r(\gamma_n)/[1 + \sigma(\gamma - \gamma_n)]$ still suffers from a violation of our condition (c).

[20] It is directly related to the radius of the pseudosphere, see also: reference 11, p. 13; H. von Helmholtz, "On the origin and significance of geometrical axioms," *Popular Lectures on Scientific Subjects*, translated by E. Atkinson (Longmans, Green and Company, London, 1881), pp. 27–72. We employ the spherical representation, such that, for $D = \infty$, $\xi^2 + \eta^2 + \zeta^2 = r^2 = [2/(-K)^{\frac{1}{2}}]^2$, where $2/(-K)^{\frac{1}{2}}$ is hyperbolic infinity (by definition) and $r_\infty = 2$ is the visual distance (in r units) to physical infinity (by choice). K may presumably be related to visual skills, in that when $K = -1$, hyperbolic infinity and the geodesic distance to physical infinity are equal; that is, one would have the actual phenomenal impression of "infinite" distance. Hence, Luneburg assumed that the closer K was to -1, the closer would be the correspondence between visual experience *at very great distances* and physical reality. However, this assumption is more of a mathematical hunch than a serious psychophysical postulate. At present, as there is no evidence on it, one way or the other, it remains highly speculative. Moreover, no special significance whatever is to be accorded the possible magnitudes of K, i.e., from $+1.0$ to -1.0. This range is simply a result of the choice (v.s.) of $r_\infty = 2$. Had we chosen $r_\infty = 20$, then K would have ranged between $+0.1$ and -0.1; if we had let $r_\infty = 1$, then K would have ranged between $+2.0$ and -2.0; this is an arbitrary point, a choice of unit size.

[21] T. Shipley, J. Opt. Soc. Am. 47, 804 (1957).

[22] F. Hillebrand, (a) Denkschr. math. naturw. Kl.d. Akd. Wien 72, 255–307 (1903). (b) *Lehre von den Gesichtsempfindungen* (J. Springer, Vienna, 1929).

judgment; no monocular cues (as they are traditionally enumerated, e.g., reference 16) can be allowed to enter. This distinction is very subtle, much more so than the traditional approach would imply.[8] However we can reinforce the distinction to some extent by ascribing it more particularly to the physiological significance of convergence, in whatever manner it may be found (or supposed) to operate.

The D distances are thus radial visual distances from the self to the points of regard; these distances emerge from the ego-center in the manner of a sunburst. Once D is known, for a given observer, we can describe the subjective visual magnitudes of all visible distances by means of the $D(\gamma)$ function. Both the differential and the finite forms of the metric may generally be used for this purpose.[1] However, for the mathematization of the alley experiments, we need only make use of the triangle equalities; these are also readily available.

Moreover, we shall find it convenient, at times, not to talk of D directly, but of some function of D, namely $r = f(D)$. The only necessary requirement of this function is that D and r be in a one-to-one correspondence. In order that we may simply represent the geodesic distance, D, in a flat diagram [e.g., Fig. 2(A)], we have the additional requirement that the r distance must be an Euclidean straight line. This is achieved, once the sign of the curvature of the metric has been agreed upon, or has been independently ascertained, by any one of the following expressions:

$$r = \frac{2}{(-K)^{\frac{1}{2}}} \tanh(-K)^{\frac{1}{2}} D/2, \quad K < 0 \qquad (3.1)$$

$$r = D \qquad\qquad\qquad K = 0 \qquad (3.2)$$

$$r = \frac{2}{(-K)^{\frac{1}{2}}} \tan(-K)^{\frac{1}{2}} D/2, \quad K > 0. \qquad (3.3)$$

Whatever properties we ascribe to the visual geodesic distance (as in Sec. C, above) must, *mutatis mutandis*, be reflected in the representational Euclidean distance.

E. THE BASIC ALLEY EXPERIMENT

As we have said, the alley experiment is the crucial experiment upon which the choice of metric depends. In fact, it is the *only* experiment dealing directly with the geodesics of the space.[23] All the other experiments on this problem; i.e., the three- and four-point studies and the frontal "horopter" measurements, are more directly concerned with the transformation equation, and the independence of form from location, than with the

[23] It seems to us that von Schelling [J. Opt. Soc. Am. 46, 309–315 (1956)] perhaps misunderstands this point, when he suggests that the alley experiment is inappropriate because the eyes are permitted to move. The very purpose of the geometry, from a physiological point of view, is to permit the mathematization of the visual space of *moving eyes*. To replace the motion of the eyes by the rotation of the head is too awkard a restriction for us to accept. Moreover, in the suggested case of an affine space (which visual space may well turn out to be), the geodesics must, *a fortiori*, still be elaborated by moving eyes!
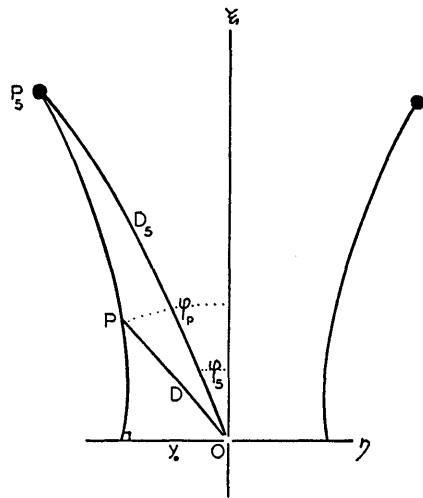

Fig. 3. A schematic representation of the parallel alleys in binocular visual space, embodying Luneburg's choice of the geodesic as normal to the η axis.

fundamental metric structure. However, the alley experiment also permits a direct investigation of both these issues; because of this dual strength, it is of particular importance in this area and warrants the most careful examination. Therefore, we present the arguments in very great detail.

Confining ourselves entirely to the triangle equalities, we can develop the transformation equation, $D(\gamma)$, from the alley experiment in the following manner. Figure 3 is a schematic representation of the parallel alleys as they exist in visual space. It may be seen that they form right angles with the apparent medial η axis of visual space. Luneburg suggested that the nature of the subject's task permits this assumption.

It is quite reasonable to assume at the outset that the parallel alleys are straight lines (geodesics) in visual space; after all, the subject affirms that they "are straight," this is his job. Without some other restriction, however, the parallel alleys may still be any one of an infinite number of straight lines. The right-angle assumption of Fig. 3 is thus equivalent to the specification of a single geodesic; it is Luneburg's choice and we shall accept it for the present. All our subsequent findings will reflect this assumption in one way or another.[24] Postulating, then, that the hyperbolic-right-triangle relations hold in visual space, we may cite the intrinsic right-triangle identity

$$\cos A = \tanh kb / \tanh kc, \qquad (4.1)$$

where $k = (-K)^{\frac{1}{2}}$ and K is the Gauss curvature of the space (expressed in units of reciprocal visual length), A is any angle (expressed in radians), and b and c are adjacent sides (expressed in units of visual length).[25]

[24] We shall present evidence later on (see reference 21) to show that this assumption is not correct in many instances.

[25] Note that this identity lies in visual space, where the geodesic distance, D, applies, not the representative distance, r. Hence:
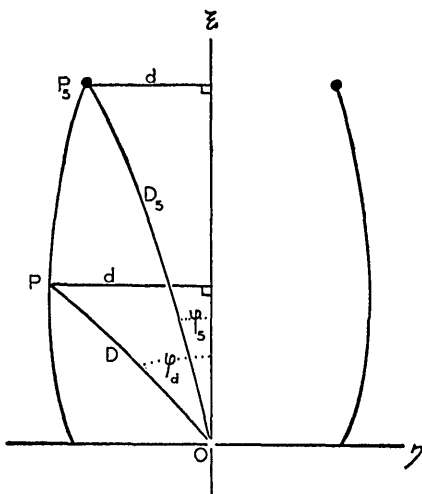
Fig. 4. A schematic representation of the distance alleys in binocular visual space, embodying Luneburg's choice of the geodesic as normal to the ξ axis.

Consider, now, the variable point P in visual space. It is seen that

$$\cos(\pi/2 - \varphi_p) = \sin\varphi_p = \tanh k y_0 / \tanh kD. \quad (5.1)$$

The same relationship holds for the fixed far point, $P_n(Q_n)$ (in our case, P_5, since $n = 5$ for the five light pairs used), through which the curve must pass

$$\sin\varphi_n = \tanh k y_0 / \tanh kD_n \quad (5.2)$$

where D_n is the maximum value, $D(\gamma) = D_n(\gamma_n)$, in the space. Setting these two expressions [Eq. (5.1) and Eq. (5.2)] equal to $\tanh k y_0$ and thus to each other, we find the basic expression for the motion of any variable point P along the curve, hence the equation of the parallel alley geodesic (entirely in visual coordinates):

$$\sin\varphi_p \tanh kD = \sin\varphi_n \tanh kD_n = \tanh k y_0. \quad (6)$$

Turning now to Fig. 4, we have a schematic representation of the distance alleys as they exist in visual space. The visual widths form right angles with the apparent median ξ axis of visual space. It is held that the nature of the subject's task permits this assumption.[26] Let us again assume that the hyperbolic-right-triangle relations hold intrinsically. We may cite, then, a second right-triangle identity

$$\sin A = \sinh ka / \sinh kc. \quad (7.1)$$

_____

$b = D_b$ and $c = D_c$ in Eq. (4.1); our succeeding discussion refers directly to D, and only indirectly to r. However, as we have said, all the general properties of $D(\gamma)$ must also hold for $r(\gamma)$, and vice versa.

[26] This also specifies one geodesic for the distance alleys, out of many possible ones, but it is much less tenuous than the choice for the parallel alleys. Two points sufficiently close together, in any Riemannian space, *always* determine a straight line; hence we know the alley width to be a geodesic. Moreover, since the alley widths are approximately symmetrical to the straight-ahead line of sight, the assumption that they are normal to this line of sight is quite reasonable.

Consider, now, the variable point P in visual space. It follows that

$$\sin\varphi_d = \sinh kd / \sinh kD. \quad (8.1)$$

The same relationship holds for the fixed far point, $P_n(Q_n)$, through which the alley must pass

$$\sin\varphi_n = \sinh kd / \sinh kD_n. \quad (8.2)$$

Setting these two expressions equal to $\sinh kd$ and thus to each other, we find the basic expression for the motion of any variable point P along the curve, hence the intrinsic equation of the distance-alley curve[27]:

$$\sin\varphi_d \sinh kD = \sin\varphi_n \sinh kD_n = \sinh kd. \quad (9)$$

Now, the far point, $P_n(Q_n)$, is common to both alleys, hence angle φ_n is identical in both curves, so that we may set the two equations [Eq. (6) and Eq. (9)] equal to each other

$$\sin\varphi_n = \sin\varphi_d \sinh kD / \sinh kD_n$$

$$= \sin\varphi_p \tanh kD / \tanh kD_n \quad (10.1)$$

from which it follows, by transposition and substitution, that

$$\sin\varphi_p / \sin\varphi_d = \cosh kD / \cosh kD_n. \quad (11.1)$$

Since k is constant, and D is always smaller than D_n, and the hyperbolic cosine varies *directly* with D, we must conclude that

$$\sin\varphi_p / \sin\varphi_d < 1. \quad (12.1)$$

This may be phrased as follows: *given the basic structure of the theory, if the version angles of the distance alley are found, by experiment, to be greater than those of the corresponding parallel alley, so that* $\sin\varphi_d$ *is greater, on the average,*[28] *than* $\sin\varphi_p$ *then the hyperbolic-right-triangle relations would hold in visual space.* (In talking of this finding, independently of the theory, we shall call it a "negative alley-discrepancy.") In this point lies the crucial test of the alley experiments: is it true that, on the average, the points of the distance alleys are wider than those of the corresponding parallel alleys?

_____

[27] Note that this is not the equation of a geodesic. The geodesic associated with the distance alley is the width of the alleys, not the length. Whereas, Eq. (6), of the parallel alley, is a geodesic equation; there, the lengths are the geodesics, not the widths. This distinction is essential. Among other things, the relationship of the distance alleys to size constancy is based upon it.

[28] We shall discuss the problem of the choice of the average later on (reference 21). At present, it is sufficient to note that an average of some sort must be taken. The errors involved in these experimental procedures are quite large, and the angles are, of necessity, quite small. It is certain that the present mathematization can only be approximate; since geometry is a highly rarified symbolism, any sensible reality must deviate from it to some extent. This point is important, because several people seemed to have approached this problem with an "either-or" attitude: either visual space is a metric space, or it is not; either it is a hyperbolic space, or it is not; and so on. This is a modern form of the classical problem of the excluded middle (a is either a, or it is not a), and is part and parcel of the so-called Aristotelian manner of thought. But, it must be recognized, psychological phenomena are largely ones of probability, and lie right in the center of the the no longer excludable middle.

A schematic diagram of this, in the two-dimensional hyperbolic plane, is shown in Fig. 5. To the extent that the complete assumptive structure herein described can be taken as valid, the close resemblance between this schema and Blumenfeld's empirical findings (Fig. 1) can be taken to indicate that the hyperbolic metric is the choice of favor. However, it must be emphasized, that this finding in itself does not in any way defend or attack the basic underlying structure. To do this requires entirely different methods.[21]

If, instead of the hyperbolic identities, Eq. (4.1) and Eq. (7.1), we had chosen the corresponding Euclidean relationships, our problem would be mathematically trivial, because $\sin \varphi_p$ would exactly equal $\sin \varphi_d$ and the ratio of the mean widths of the two alleys would be unity, i.e., the alleys would be identical. Some Euclidean cases have been reported in the literature; their meaning has not yet been elucidated (reference 21, Fig. 3; also reference 12).

Retaining the basic assumptive structure intact, changing only the sign of the metric, the choice of the corresponding nontrivial elliptical relationships entails the following intrinsic elliptical-right-triangle identities:

$$\cos A = \tan kb / \tan kc \qquad (4.2)$$

and

$$\sin A = \sin ka / \sin kc. \qquad (7.2)$$

In this case, we would replace Eq. (10.1) with

$$\sin \varphi_n = \sin \varphi_d \sin kD / \sin kD_n$$
$$= \sin \varphi_p \tan kD / \tan kD_n \quad (10.2)$$

and Eq. (11.1) with

$$\sin \varphi_p / \sin \varphi_d = \cos kD / \cos kD_n. \qquad (11.2)$$

FIG. 6. Assuming Luneburg's choice of geodesics, the theoretical positions of the two alleys are shown, schematically, in an elliptical plane: – – – – – parallel alleys; ———— distance alleys.

Then, with k being constant as before and D always smaller than D_n, and since the circular cosine varies *inversely* with D, it would occur that

$$\sin \varphi_p / \sin \varphi_d > 1. \qquad (12.2)$$

This may be phrased as follows: *given the basic structure of the theory, if the version angles of the distance alleys are found, by experiment, to be less than those of the corresponding parallel alleys, so that $\sin \varphi_d$ is, on the average, less than $\sin \varphi_p$, then the elliptical-right-triangle relations would hold in visual space.* (In talking of this finding, independently of the theory, we shall call it a "positive alley-discrepancy.") A schematic diagram of this, in the two-dimensional elliptical plane, is shown in Fig. 6. With full cognizance of the logical structure herein developed, the marked discrepancy between this schema and Blumenfeld's empirical findings (Fig. 1) allows us to reject the elliptical metric.[29]

We have now only to derive the hyperbolic expression for D_n; once we have satisfied the conditions of Eq. (12.1), we shall then be able to find all the points of the $D(\gamma)$ curve, according to either or both Eq. (6) and Eq. (9). We may put, as above, $\sin \varphi_p / \sin \varphi_d$

$$= \cosh kD / \cosh kD_n$$
$$= \sinh kD \tanh kD_n / \tanh kD \sinh kD_n. \quad (11.1)$$

By transposing and squaring both sides, we have

$$\cosh^2 kD = \cosh^2 kD_n \sin^2 \varphi_p / \sin^2 \varphi_d. \qquad (13)$$

Then we may transpose and square certain other terms,

[29] Hardy *et al.*[12] and P. Squires [A.M.A. Arch. Ophthalmol. **56**, 288–297 (1956)] have reported some "elliptical" cases. The significance of these is unclear, because of possible various interpretations of the instructions, in the former case, and of the "photopic" observational conditions, in the latter. von Schelling[23] believes that such findings are based on a misunderstanding.
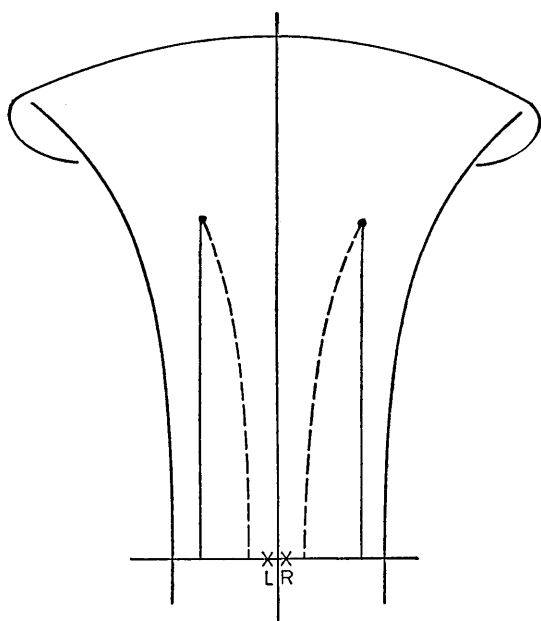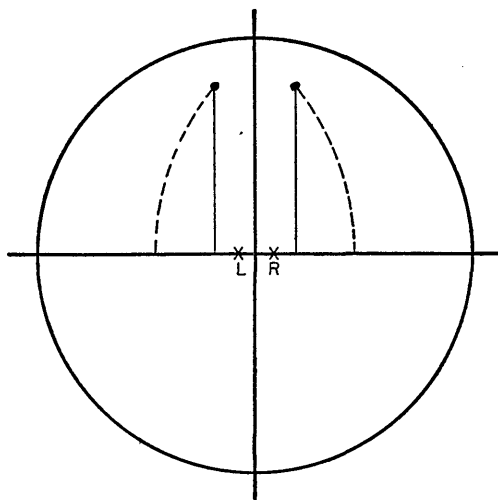

FIG. 5. Assuming Luneburg's choice of geodesics, the theoretical positions of the two alleys are shown, schematically, in a hyperbolic plane: – – – – – parallel alleys; ———— distance alleys.

from Eq. (10.1), to give us

$$\sinh^2 kD = \sinh^2 kD_n \, \sin^2\varphi_n / \sin^2\varphi_d. \quad (14)$$

Now we recall the following identity

$$\cosh^2 ka = 1 + \sinh^2 ka$$

so that we may put

$$\sinh^2 kD_n \, \sin^2\varphi_n / \sin^2\varphi_d = \frac{\cosh^2 kD_n \, \sin^2\varphi_p}{\sin^2\varphi_d} - 1. \quad (15)$$

From this it follows, by substitution and cancellation, that

$$(\cosh^2 kD_n - 1)\, \sin^2\varphi_n = \cosh^2 kD_n \, \sin^2\varphi_p - \sin^2\varphi_d \quad (16.1)$$

or

$$\cosh^2 kD_n \, \sin^2\varphi_n - \sin^2\varphi_n$$
$$= \cosh^2 kD_n \, \sin^2\varphi_p - \sin^2\varphi_d, \quad (16.2)$$

and, solving the expression for $\cosh^2 kD_n$, we have, finally,

$$\cosh^2 kD_n = \frac{(\sin^2\varphi_d - \sin^2\varphi_n)}{(\sin^2\varphi_p - \sin^2\varphi_n)}. \quad (17.1)$$

With this equation, we have derived, in detail, all the relevant equations for the complete mathematization of the alley experiment, with due regard for the necessary empirical limitations on γ. This includes the specialized expression, Eq. (17.1), for the limiting value $[D_n(Q_n) = D_n(\gamma_n > 0)]$ of the fixed far point. When we substitute this value into the equations for the two alleys, Eq. (6) and Eq. (9), we will be able to find the other empirical values of $D(\gamma)$. This determines the $D(\gamma)$ function in its entirety.[30]

It is important to mention, at this point, that in Eq. (6) and Eq. (9) we have a very subtle check on the assumptive structure of the theory—particularly upon the right-angle assumptions of Fig. 3 and Fig. 4. [Recall that Eq. (12.1) and Eq. (12.2) were our first check points, assuming the validity of this structure.] If the $D(\gamma)$ function is to "characterize" the visual space of a given individual,[31] then Eq. (6) and Eq. (9) must both give the same $D(\gamma)$ function, within the experimental error; the right-angle assumptions must hold rigidly for both kinds of alleys. We shall report several cases later on where this is not so.

F. THE NORMALIZATION

One point remains to be discussed. In Eq. (17.1), the right-hand side contains the empirical findings, the left-hand side contains the theoretical constants which we are trying to determine. But there are two of them, k and D_n. Both cannot be determined simultaneously from a single equation. The difficulty can only be

resolved by *a priori* means. The problem is that of the choice of unit, for which empirical solutions are not possible.[32]

Only three procedures exist: either to assume a value for k (hence K) and find all the empirical magnitudes in terms of units of reciprocal visual length (reference 21, Fig. 10 through Fig. 13); or to assume a value for D (any D, it matters not which) and express k (hence K) and all the empirical magnitudes in terms of units of visual length; or to assume a value for the product of k (hence K) and some D and express all the empirical magnitudes as pure numbers (reference 21, Fig. 7 and Fig. 8).

That only these three *a priori* solutions exist can be shown substantially as follows. The description of any measuring process necessarily involves (1) a unit of magnitude (in the present case, one particular radial visual length), and (2) an operation (in the present case, the estimation of some particular radial visual length), and finally, (3) some magnitude to be described (in the present case, any radial visual distance). We can choose the unit, and describe the distance in terms of the number of operations; we can choose the number of operations (in units of reciprocal magnitude or length), and describe the distance in terms of the size of the unit; or we can, boldly, take the first distance to be measured as the unit. For this, we typically choose a very singular distance, e.g., the binocular horizon or the experimental far point, and describe all other distances in terms of this. In this latter case, we have, in effect, taken the product of *"size of unit times number of operations"* as unit, and expressed all other magnitudes in terms of this dimensionless unmeasured quantity.

There appear to be no reasons, other than purely pragmatic ones, why we must, or should, choose one or the other of these normalization procedures. In a Riemannian space of constant curvature, there can be no difference, eventually, between the three choices. Hardy *et al.*[12] and Zajaczkowska,[13] for example, have both normalized according to $D(\gamma = 0)$, as suggested by Luneburg; it is equally possible to normalize with respect to the far point, $D_n(\gamma_n)$. And, as we have already noted, it is not easy to get a meaningful, or credulous, response to: "How far away are the stars?" (see Gilinsky, reference 10).

Independently of the normalization, however, Luneburg[1c] (p. 637) has also shown that our Eq. (17.1) may be expressed as:

$$\cosh^2 kD_n = \frac{\sin^2\varphi_d - \sin^2\varphi_n}{\sin^2\varphi_p - \sin^2\varphi_n} = \left[\frac{1 - \dfrac{K}{4}D_n^2}{1 + \dfrac{K}{4}D_n^2}\right]^2 = \frac{Y_{0d}}{Y_{0p}}, \quad (17.2)$$

---

[30] The transformation equations from the actually recorded Cartesian coordinates to the bipolar coordinates are as follows: $\tan\theta = z/x$; $\tan\phi = (y \cos\theta)/x$; $\gamma = (\frac{1}{2}PD \, \cos^2\phi \, \cos\theta)/x$.

[31] See L. H. Hardy, A.M.A. Arch. Ophthalmol. **42**, 551–561 (1949); also references 11 and 13.

[32] Recall that one-half the interpupillary distance, taken in millimeters, is our corresponding choice of unit in physical space [see Fig. 2(B)].

or, in terms of K and D_n:

$$KD_n = \frac{4(Y_{0p}/Y_{0d})^{\frac{1}{2}} - 1}{(Y_{0p}/Y_{0d})^{\frac{1}{2}} + 1}, \qquad (17.3)$$

where Y_{0p} and Y_{0d} are the intersections, with the η axis, of tangent lines to the respective parallel- and distance-alley curves, taken at the common far point of the alleys. There can be no basic objection to this expression as it stands, but it does have an inherent weakness, and has encouraged some work which we believe should be critically reviewed.

Serious comments on these previous studies center about the subsequent modification of this equation:

$$K = \frac{e^{2\sigma\gamma}(Y_{0p}/Y_{0d})^{\frac{1}{2}} - 1}{(Y_{0p}/Y_{0d})^{\frac{1}{2}} + 1}. \qquad (18.1)$$

Here, the $r(\gamma)$ function, $r(\gamma) = 2e^{-\sigma\gamma}$, is assumed to be as Luneburg gave it; *but this very function is*, we believe, *the most important point yet at issue!* And we shall show[21] that other functions will do as well, and perhaps better (see also reference 11). Zajaczkowska[13] did some preliminary experiments to justify her choice of the Luneburg function, but these are not reported in detail. In any case, it is doubtful that she has shown this to be the actual form of the function in preference to any number of other equations, but merely that this does bear a reasonably close approximation to the empirical data—which is a somewhat different thing. However, be this as it may, her further determination of σ from the three-point experiment, entirely independent of the alleys, is an ideal procedure. The integration of results from independent experiments is, in a certain sense, the eventual goal of the theory.

This cannot be said, however, of the procedure used by Hardy et al.,[12] where both σ and K were determined from the single alley experiment, and where $r(\gamma)$ was assumed without the benefit of preliminary studies.

Moreover, Eq. (18.1) encourages, by its enticing simplicity, an inherent defect which has affected both these studies, and which is fairly serious from a psychophysical point of view. Only one experimental point from each alley curve is actually used in the determination of the intercept values, Y_{0p} and Y_{0d}. The tangent lines were computed as passing through the fixed far point, Q_n, *which is not an experimentally variable point in itself*, and through the single variable point, Q_{n-1}. All the other variable points, between the observer and point Q_{n-1}, and these are phenomenologically essential points, were ignored. Such a procedure wastes what may be extremely valuable data. It must markedly decrease the reliability of the results, and may actually give a false impression. To defend it on the basis that it avoids

the practical difficulties of a polynomial curve fitting procedure is certainly possible, but we feel that such difficulties must be faced squarely, particularly because the delicacy of the present experimental environment warrants the greatest possible mathematical sensitivity. Moreover, any genuine psychophysical function must be based on integrating or averaging procedures,[28] albeit sometimes very complex, else it would lack definitive accord over the complete phenomenal dimension.

In view of these considerations, we have normalized in terms of K (see also reference 11). K is taken as -1 such that $k=1$, hence $D_n(\gamma_n)$ and $D(\gamma)$ are expressed in units of $1/k$. This simplifies Eq. (17.1), such that

$$\cosh^2 D_n = (\sin^2\varphi_d - \sin^2\varphi_n)/(\sin^2\varphi_p - \sin^2\varphi_n). \quad (18.2)$$

The accumulation of evidence up to the present time does indicate that the alley-discrepancy, though sometimes apparently zero, is most often negative.[11,13,29,33] Given the basic assumptive structure of Luneburg,[1] to describe binocular visual space as a hyperbolic space of constant negative curvature does seem to be reasonably well justified, at this point in our discussion. Moreover, the assumption that K is constant for each individual observer, at least in the visual direction of "straight-ahead at eye-level" (we shall use the more direct German phrase here: the *Gerade-Voraus*), is fairly compelling on its face, in view, among other things, of the shape-constancies.[34] Thus, we choose to work with Eq. (18.2), and not with Eq. (18.1).

SPECIAL ACKNOWLEDGMENTS

[33] T. Shipley (unpublished thesis, New York University, 1955) has also reported some reversals in the parallel-distance-alley discrepancies. These were for one observer, in regions of visual space at various version angles from the straight-ahead direction. However, due to the uncertainties in rotations of axes, resulting from our present dissatisfaction with Luneburg's choice of the parallel geodesic (reference 21), the significance of these findings is not yet clear. We shall, therefore, not discuss them in the present context; the raw data are readily available from the author. It should be added, however, that the other observer gave consistent alley discrepancies in all regions of visual space—these also cannot yet be interpreted.

[34] T. Shipley (reference 33) has found individual differences in the constancy of K, on the basis of variable alley discrepancies, and notes its relation to aniseikonic distortions. However, for normal observers, or fully corrected observers, the assumption of constancy may perhaps stand.