Journal of Mathematical Biology 1, 97—131 (1974)

Differential Geometry and Color Perception

H. L. Resnikoff, Houston

Received February 11, 1974

Summary

Combined with well known facts about the affine structure of color space $\mathscr{P}$, a local homogeneity hypothesis implies that $\mathscr{P} = G/K$ is equivalent to either $\boldsymbol{R}^+ \times \boldsymbol{R}^+ \times \boldsymbol{R}^+$ or $\boldsymbol{R}^+ \times SL(2, \boldsymbol{R})/SO(2)$. The G-invariant metrics yield in the first case Stiles' generalization of Helmholtz' color metric, in the second a new color metric with respect to which $\mathscr{P}$ is not isometric to a Euclidean space. The paper contains an extensive historical introduction.

1. Historical Introduction

The current theories of color perception provide an interesting and important example of the application of mathematics to the description and elucidation of physical and psychophysical phenomena. Although many great thinkers have held that an analytical or mathematical treatment of the subject is impossible or even undesirable, they have gradually deserted the field so that today and indeed throughout the past 50 years it has been generally recognized that a theory of color perception must be, both in form and content, a mathematical theory.

Plato is no doubt the most prominent representative of the earliest opinion of our subject, that the mystery posed by the mutability of color transcends the possibility of human comprehension. In *Timaeus* he wrote (cf. MacAdam 1970)

> The law of proportion according to which the several colors are formed, even if a man knew he would be foolish in telling, for he could not give any necessary reason, nor indeed any tolerable or probable explanation of them.

and concerning the mixtures of colors,

> ... He, however, who should attempt to verify all this by experiment would forget the difference of the human and the divine nature. For God only has the knowledge and also the power which are able to combine many things into one and again resolve the one into many. But no man either is or will be able to accomplish either the one or the other operation.

Apart from its presumptive aspects, this passage is a concise and accurate description of the celebrated experiments performed more than 2000 years later by Isaac Newton, experiments which mark the inception of the modern and fundamentally mathematical view of color perception.

Plato's strictures obviously could not lead anywhere, and they play no further role in the development of the understanding of the perception of color. Aristotle's opinions, however, continued to influence thinkers until the middle of the nineteenth century, and are responsible in part for some of the most disgraceful and counterproductive diatribes of scholars (we cannot say "scholarly diatribes") ever to have appeared in print. Thus one finds the influential French Jesuit Louis Bertrand Castel accusing Newton of bad faith and stupidity, as well as faulty observation (Schier 1941), and, in the next century, Goethe (Goethe 1810), who, as Helmholtz remarks, was usually even-tempered and courtierlike, was stirred to refer to Newton as "incredibly impudent", a "Cossack Hetman", to call his work "mere twaddle", "ludicrous explanation", "admirable for children in a go-cart", and, ultimately, to attack his veracity: "but I see nothing will do but lying, and plenty of it". Evidently the study of color was then pursued with greater passion than is now considered necessary.

The Aristotelian theory which called forth such vigorous defense and extensive explication is set forth in *On the Soul, Sense and the Sensible*, and the *Meteorologica* (cf. MacAdam 1970). In the first work Aristotle concludes the necessity of a medium between the eye and an object it views; in the second he asserts that

> ... as vision would be impossible without light [between the object and the eye], so also would it be impossible if there were no light inside [the eye].

and goes on to propose the theory, adopted by Goethe much later, that color is a mixture of white and black. We need not review here the typical train of misleading deductions founded on straightforward observations which led him and his numerous followers to this unfortunate conclusion. The third work concludes, insofar as color is concerned, that the rainbow exhibits precisely the three colors red, green, and violet, in that order. The

> appearance of yellow is due to contrast, for the red is whitened by its juxtaposition with green.

These passages in *Meteorologica* foreshadow in a curious way the three color theory of Thomas Young (*vide infra*) and focus, perhaps for the first time, on the mutability of perceived colors as a function of the colors of their surroundings.

We shall return to this issue in the mathematical portion of the present work.

It is important to recognize that not only was Aristotle's theory (that color is a mixture of white and black) wrong, but it never led, and never could lead, to any quantitative description of the results of color mixing experiments; his theory, and those based upon it, never rose above the stage of verbal classificatory descriptions.

The first theory which associated distinct colors with distinct and quantifiable physical states appears to be due to Descartes. In assessing his work it is well to bear in mind that he lived in an age of great intellectual ferment due to events which greatly altered the prevalent conception of the structure of physical reality. The discovery of the Americas and their pagan civilizations, the circumnavigation of the globe, the invention of the telescope and consequent observation of the miniature solar system constituted by Jupiter and its moons all conspired to destroy scholasticism and to demand the erection of a new world order. Descartes undertook this task. When faced with the particular problem of accounting for the mutual interactions of bodies not in contact with each other, such as magnets or the moon and earth as evidenced by the tides, he was forced, in order to remain within the realm of experiental certainty, to conclude that there cannot be action at a distance, but only the well understood contact forces of pressure and impact. He was consequently led to postulate the existence of a pervasive medium — the aether — by means of whose mechanical properties forces were transmitted from one point to another. The constituent particles of the aether were presumed to be continually in motion, but, due to the latter's pervasiveness, the motion of one particle entailed the motion of a chain of others. When applied to the rainbow with its "diversities of color and light", this theory suggests that

... the various colors are connected with different rotatory velocities of the globules, the particles which rotate most rapidly giving the sensation of red, the slower ones of yellow and the slowest of green and blue — the order of the colors being taken from the rainbow. The assertion of the dependence of color on periodic time is a curious foreshadowing of a great discovery which was not fully established until much later (Whittaker 1951; *cf.* Descartes 1638).

We note that according to this theory, each perceived color must correspond to a state of rotation and that distinct hues correspond to distinct angular velocities. Thus there is, according to Descartes, a bijective correspondence between a certain set of rotation states and the set of perceived hues. That this claim is false could readily have been determined by Descartes from elementary observations, but he and many others as well were misled by the presumption that the human senses faithfully transform distinct physical stimuli into distinct sensory responses. The development of the theory of color perception was hindered until the 19th century when it was realized that such a theory must necessarily be distinct from a theory of the physical properties of light whose correlates are the perceived colors.

Although Descartes' theory admits quantification, nothing seems to have been done to determine the rotatory velocities which were presumed to correspond to particular colors, nor was any attempt made to provide a more precise description of the physical distinctions which give rise to the subjective discrimination between hue and saturation. Thus, according to the theory as stated above, one must presume a bijective mapping of the two dimensional continuum of colors of varying hue and saturation onto an interval of the real line whose points are identified with the various rotational velocities ("brightness" is readily associated with the number of light quanta, and need not concern us here). We know today how necessarily complicated such a bijection is, and, armed with this knowledge, we might prefer to forgo bijectivity in favor of avoiding the introduction of so complicated a function. This simply amounts to the recognition that the visual sensory system does not provide a faithful representation of physical reality.

In elaboration of his theory of color, Descartes supposed that the rotatory motion perceived as color was acquired through successive refraction of the light impulse as it was transmitted through the aether by static pressure. This notion, and the entire conception of color to which it corresponds, was attacked by Robert Hooke in 1665 in his *Micrographia*. By his study of the colors of thin plates, that is, the colors seen when light falls on a thin layer of air bounded by two parallel transparent plates, Hooke demolished Descartes' theory and then turned to the presentation of his own views. Hooke's theory is transitional between Descartes' theory and the fully developed wave theory of light. He concludes the light "must be a *Vibrative* motion" (Hooke 1665), propogated in waves, and that the origin of color lies in the deflection of the wave front from the perpendicular to the direction of motion of the light pulse due to characteristic refractions and reflections. In this instance, the light pulses may be supposed

> ... *oblique* to their progression, and consequently each Ray to have potentially *superinduc'd* two properties, or colours, *viz*. a *Red* on the one side, and a *Blue* on the other, which not withstanding are never actually manifest, but when this or that Ray has the one or the other side of it bordering on a dark or unmov'd *medium*, therefore as soon as those Rays are entred into the eye, and so have one side of them each bordering on a dark part of the humours of the eye, they will each of them actually exhibit some colour.
>
> .. we may collect these short definitions of Colours: *That Blue is an impression on the Retina of an oblique and confus'd pulse of light, whose weakest part precedes, and whose strongest follows.* (Hooke 1665, p. 64).

A similar definition of red follows, in which the roles of the "weak" and "strong" part are interchanged. Hooke recognizes that according to this theory there are "but two Colours" and argues (Hooke 1665, p. 67)

> that the *Phantasm* of Colour is caus'd by the sensation of the *oblique* or uneven pulse of light which is capable of no more varieties than two that arise from the two sides of the *oblique* pulse, though each of those be capable of infinite gradations or degrees (each of them beginning from *white*, and ending the one in the deepest *Scarlet* or *Yellow*, the other in the deepest *Blue*).

The colors are thereby identified with the real line as a geometrical continuum, and may be parametized, for instance, by the tangent of the angle of obliquity of the incident wave front, with white corresponding to normal incidence. Although Hooke's theory of color perception is false in most respects, it does mark an important milestone along the path toward recognition that the "Phantasm" of colors is in part subjective.

Hooke's theory was short-lived, for in 1671 Newton published the results of his elegant and decisive experiments which showed that by means of a transparent prism, a beam of light could be decomposed into primitive "spectral" constituents which resisted further decomposition, that superposition of the constituents of a light resulted in the reconstitution of the original light, and that there are perceived colors — specifically, white — which do not occur as spectral constituents. From our viewpoint, Newton discovered a means for determining the absolute value of the Fourier transform of a visible light signal, and, by means of the principle of superposition, established that the space of physical light stimuli can be identified with the subset of non-negative functions in the Hilbert space of square integrable functions whose domain is a closed interval of real numbers. In particular, the space of physical lights is an open cone in an infinite dimensional vector space, a result which could not be more clearly in contradiction with Aristotle's theory of the constitution of color from black and white, or with

Hooke's more profound mechanically conceived two-color theory. Hooke's subsequent attack on Newton is no doubt responsible in part for the delay before the appearance of the latter's *Opticks* (Newton 1704).

Newton's color circle, which will be of importance in what follows, makes its first appearance in the solution of Proposition VI Problem II of the *Opticks*, to wit:

In a mixture of primary colors, the quantity and quality of each being given, to know the color of the compound.

By "compound" is meant "superposition". The problem asks for that primary, *i.e.* spectral, hue which *perceptually* matches the superposition, and also for a quantitative measure of

its distance from whiteness,

that is, of the *perceptual* saturation of the resultant color. Newton's solution ranges the spectral hues along the circumference of a circle in their natural order of occurrence in the spectrum with the red and indigo extremes conceived as gradually passing into a common violet. White is placed at the circle's center. If various spectral hues be given, Newton constructs circles whose centers are the given hues and whose areas "are proportional to the number of rays of each color in the given mixture". The centroid of the resulting configuration specifies a point in the color circle whose distance from the origin (white) is a measure of the saturation of the superposed resultant; its spectral hue is determined by the intersection with the circle of spectral hues of the radius which passes through the centroid. The quantity of the resultant light is the sum of the areas of the circles corresponding to its constituents, so that a light depends on three variables: *quantity*, *hue*, and *saturation*. Newton's assertions mean that the set of lights forms a three dimensional Euclidean manifold which can be identified with the set of pairs $\{(\alpha, x)\}$ where x denotes a point in Newton's color circle and α is a positive real number which corresponds to the quantity of light; this manifold can be equivalently conceived as a cone in three dimensional Euclidean space.

It is an immediate consequence of this procedure that infinitely many distinct mixtures of light must be perceived as identical and therefore that a theory of color perception cannot be simply a theory of the physical properties of light; the nature of the visual sensory system itself must play an essential role.

Newton was aware that his geometrical rule is but an approximation; it was he said,

. . . accurate enough for practice, though not mathematically accurate (Newton 1704).

Although Helmholtz asserts that the comparison between the spectrum colors and musical notes was first suggested by Newton (Helmholtz 1866), according to Ernst Mach (Mach 1913) an analogy of this type had already been drawn by the astronomer Ptolmey (*circa* + 150), and a parallelism between color and the pitch of sound appeared probable to Descartes as a consequence of his dynamical theory. But Newton was certainly the first to attempt to detail the hypothesized correspondence. His color wheel is partitioned into seven sectors derived from the widths of the colored areas in the spectrum of a glass prism which are meant to correspond to the notes of the diatonic scale. As the spectrum is a continuum, this partition is arbitrary and does not have any intrinsic meaning. Indeed, if one broadly classifies the subject matter of mathematical theories as *geometrical* in the sense of the study of continuous manifolds, or *arithmetical* in the sense of the study of discrete systems, then in general a specific problem will partake of the characteristics of one or the other but not both of these classes. Now since the time of Pythagoras the theory of musical harmony has been known to lie distinctly in the arithmetical class; therefore, Newton's effort to identify his strikingly successful geometrical theory of color with the arithmetical theory of the diatonic scale was as strange as it was doomed to failure. Nevertheless, this idea continued to intrigue lesser minds for more than a century. Thus, Castel attempted to apply the supposed analogy to the construction of an "ocular harpsichord", intended to produce pleasing effects by a display of colors calculated to correspond to the particular notes, both sound and color to be called forth simultaneously by the play of the keys. His efforts, spanning the period 1725 to 1754, were greeted with enthusiastic anticipation by the French intelligentsia and drew financial support as well as encouragement from celebrated figures such as Montesquieu. A prototype instrument containing 60 colors was ultimately constructed but, based as it was upon an unreal analogy, it was no more than a short-lived curiosity. Spurred by Frauenhofer's measurements of the wave-length of light of various colors, elaborate efforts were once again made, particularly by Drobisch and later by Unger, to find a connection between the musical scale and the variation of color; the interested reader should consult Helmholtz. The main result of these speculations has simply been to establish that a description of the phenomena of color must be essentially geometrical, not arithmetical, and that one must seek for further substantial improvements in their mathematical description amongst the more novel and deeper aspects of geometry.

In Newton's work there is still confusion between physics and psychophysics; he supposed that the eye as an instrument measures physical differences perceived as color by means of a bijective correspondence between the set of physical input light signals and the set of sensory responses. This confusion persists in Euler, who identified the variation of physical color stimuli with frequency variation but failed to recognize that the frequency of a light wave is not determined by its subjective color (Euler 1752).

The formulation of a theory which explicitly recognized the distinction between the physical properties of light and the correlated subjective sensations it produces is the achievement of the multi-faceted Thomas Young. The physiological mechanism for converting light into the sensation of color which he proposed has been confirmed in its essential features and is the foundation upon which current theory has been erected:

As it is almost impossible to conceive each sensitive point of the retina to contain an infinite number of particles, each capable of vibrating in perfect unison with every possible undulation, it becomes necessary to suppose the number limited, for instance to the three principal colors... Each sensitive filament of the nerve may consist of three portions, one for each principal color. (Young 1801; cp. Young 1845.)

Young's theory lay dormant for 50 years until Helmholtz in Germany and James Clerk Maxwell in England rescued it from oblivion. They confirmed and strengthened its essential features in numerous ingenious experiments (Helmholtz 1852, 1866; Maxwell 1856, 1857, 1860), and thereafter the development of the theory of color perception proceeded rapidly. In 1853 the mathematician Grassmann stated the results of color mixture experiments in a form very nearly equivalent to the assertion that the set of colors is a convex cone in 3-dimensional affine space; the following year in his Habilitationsschrift "Ueber die Hypothesen, welche der Geometrie zu Grunde liegen", Riemann observed that

The positions of the objects of sense, and the colors, are probably the only familiar things whose specifications constitute multiply extended manifolds (Riemann 1854),

a pregnant remark whose implications remained hidden until Helmholtz brought them to light in 1891. Thus, although Maxwell was probably familiar with Riemann's conception of metrical geometry, in his 1857 paper "The diagram of colors", he retained the Euclidean nature of Newton's color space, identifying Newton's color circle (and his own modification of it) with a 2-dimensional Euclidean subspace corresponding to lights of a fixed brightness, and failed to realize that the measure of brightness need not be a multiple of the measure of physical light intensity. It was undoubtedly Fechner's solution of this latter problem that made the significance of Riemann's remark clear to Helmholtz. In his study of the apparent magnitude of stars, Fechner was led to conclude that the measure of perceived brightness varies as the logarithm of the incident light intensity (Fechner 1877; cp. Resnikoff). We may therefore say that the physical metric of the intensity continuum is different from the psychophysical metric of the continuum of perceived brightness, and that the logarithm function provides the means for transforming the former magnitude to the latter. In terms of Riemannian geometry, the two 1-dimensional spaces are isometric, and therefore not essentially distinct. Helmholtz (Helmholtz 1891, 1892) had the happy idea of extending this notion to the entire 3-dimensional color space, that is, of defining a Riemannian metric on the space of perceived colors which would measure the perceptual difference between any two colors. Color space with this metric need not be isometric to Euclidean space, although it is in the particular model introduced by Helmholtz.

According to observations of Abney (Abney and Festung 1886, Abney 1913) brightness is a linear function of perceived lights. Geometrically, this means that the surfaces of constant brightness are portions of planes in the 3-dimensional Euclidean space in which the color space is embedded. The surfaces of constant brightness derived from Helmholtz's metric are not planar; this disagreement with Abney led Schrödinger to undertake a re-examination of the entire affine and metrical theory of color perception and resulted in his proposal of a new metric compatible with both Abney's and Fechner's results but with respect to which color space is not isometric to Euclidean space (Schrödinger 1920). More recent experimental studies have shown that in fact brightness is not a linear function (Graham 1965), so this one time pillar of Schrödinger's theory has become one of its several serious drawbacks. Nevertheless, his paper is still of methodological importance and is a model of elegance.

In 1946 Stiles introduced an improvement of the Helmholtz metric which amounts to a change of the unit of measure of each coordinate in the underlying Euclidean space (Stiles 1946, Graham 1965, Stiles and Wyszecki 1967), but this simple modification significantly improves the accuracy of the description which the metric provides. The consequences of Stiles' metrical theory have been worked out in elaborate detail in the publications noted above and are in generally fair agreement with observations.

2. Introduction to this Work

The path from Plato to Stiles by way of Newton, Young, and Helmholtz has led to an increasing geometrization of the theory of color perception and to the introduction of Riemannian differential geometry as a basic tool for the analysis of subjective color phenomena. Nevertheless, there remains an element of arbitrariness in the selection of a color metric and there are various well known observational phenomena which play no role in the current theory.

One of the most important of these phenomena is the invariance of relative color perception. Already in 1876 Helmholtz brought attention to the paradox of the painter's ability to represent greatly different states of illumination intensity with pigments. It is evident that the painter of a picture in the classical style does not attempt to produce the same distribution of light and color that would be incident upon the eye if the original scene were viewed. Indeed, if two pictures be hung adjacent to each other in a gallery, one of a bright desert scene exhibiting various degrees of whiteness and dark shadow and the other of a moonlit night or other dark interior scene with whitish highlights, then the actual quantities of light reflected by the bright (resp., dark) parts of either picture to the eye of a beholder will be approximately the same since in both cases the same white (resp., black) pigment will have been used. Moreover, the brightest white on a picture in a gallery as ordinarily lit is perhaps but 1/40-th the brightness of that white directly lit by the sun; if viewed in the desert, the painting of the desert as lit in the gallery would appear dark grey. More remarkable still, the brightest white pigment reflects only about 100 times as much light as the darkest black whereas the sun's disk is about 80,000 times as luminous as the disk of the full moon. The artist consequently has no hope of reproducing the true light distribution of the original scene with his palette of limited hues and range of brightness, and therein lies his art. For according to Fechner's work, the perceived brightness $b(x)$ of a light of intensity x is proportional to $\log x$ and consequently the relative brightness $b(x_1) - b(x_2)$ of the lights of intensity x_1 and x_2 will be proportional to

$$\log x_1 - \log x_2 = \log\left(\frac{x_1}{x_2}\right).$$

We conclude that relative brightness is invariant under the simultaneous modification of light intensity (Cornsweet 1970)

$$x_1 \mapsto \lambda x_1,\ x_2 \mapsto \lambda x_2,\ \lambda > 0;$$

Hence, on the whole, the painter can produce what appears an equal difference for the spectator of his picture, notwithstanding the varying strength of light in the gallery, provided he gives to his colours the same *ratio* of brightness as that which actually exists. (Helmholtz 1876.)

And moreover it is this *ratio* which characterizes the state of illumination of pairs of familiar objects and thereby enables us (exclusive of extremes states of illumination for which Fechner's Law fails) to recognize both the real conditions of illumination and their artistic representations.
If X denotes a topological manifold and G a group of continuous transformations of X onto X, and if G acts transitively on X, that is, if to every pair of points x and y of X there corresponds a transformation $g \in G$ which carries x to y, thus $g(x) = y$, then X is said to be a *homogeneous space* of G. The possible degrees of intensity of illumination can be identified with the set R^+ of positive real numbers; R^+ is a topological manifold and also a group. Since $y = (\frac{y}{x})\, x = \lambda x$ for arbitrary $x, y \in R^+$, $X = R^+$ is a homogeneous space of $G = R^+$, and the remarks about brightness offered above show that relative brightness is a G-invariant function defined on X. What is more, up to selection of the unit of measure, Fechner's Law defines the unique G-invariant metric on X.

The purpose of the present paper is to generalize this construction to the entire color space. It will be shown that only two types of homogeneous space are compatible with a small number of well attested hypotheses and, what is more surprising, that a metric invariant under the action of the transitive group is essentially uniquely determined and perceptual measures are two-point invariant functions. In the one case the metric is that of Stiles, but the other case appears to be new.

Common experience provides considerable evidence that the entire color space is homogeneous with respect to some group. Thus, since not only the intensity but also the color of illumination in a daylight-lit gallery varies depending on the time of day, the cloud cover, and the season without significantly affecting the relative hues, saturation, or brightness of components of the exhibited paintings, we must conclude that their relative properties are independent of the state of illumination throughout a large region of color space which includes all normal conditions of illumination. That it is not necessary to relearn color relations when one uses tinted sunglasses or drives an automobile with a tinted windshield provides additional corroboration, and the essentially interchangeable use of incandescant and fluorescent illuminants supports this contention in still another way. But the most convincing evidence of homogeneity and its fundamental role in the visual perception process is found in an important and striking series of recent experiments (Riggs *et al.* 1953, Cornsweet 1970) concerned with fixated vision. These experiments showed that if the small involuntary continuous ("*saccadic*") motions of the eye relative to a fixed and unvarying scene are cancelled by an ingenious system of mirrors, so that the image of the visual scene is motionless on the retina, the stabilized image when first turned on, looks very sharp and clear, but then it rapidly fades out and disappears and the field looks uniformly grey. Stabilized patterns disappear within seconds, or even fractions of a second after being presented. After the image has disappeared if it is moved across the retina, . . . it will reappear and then quickly disappear again. (Cornsweet 1970.)

This is just the result that would be anticipated if color space is homogeneous, for in that case single colors are not discriminable and the stationary neural activity of each light receptor in the retina is construed as a fixed neutral color independent of the actual composition of the physical illumination of the retina. In other words, there does not exist a color which can be absolutely discriminated by the human visual system, all colors are equivalent, and only certain functions of pairs of colors presented to the retinal light receptors within sufficiently brief time intervals have any perceptual significance.

The remainder of this paper is devoted to the mathematical development of a theory of color perception based upon the presumed homogeneity of color space subject to certain limitations of the scope of the inquiry to which we now turn.

We will be concerned solely with the establishment of a phenomonological theory of color perception. It is not the purpose of this work to connect the observed absorption properties of the cone and rod pigments with geometrical properties of color space, nor will the structure of the neural processing network be considered here.

These questions (especially the former), the explicit connection between the geometrical quantities introduced here and the usual parameters of color vision research, and, not least important, the quantitative aspects of the agreement of the theory herein proposed with experimental results, will be taken up in a sequel to this work. The justification for presenting a mathematical theory of a physical phenomenon without a discussion of its quantitative agreement with observation is twofold: first, the theory does appear to be in qualitative agreement with observations, and second, it is grounded on accepted experimental results of so qualitative and general a nature that it cannot be expected to provide a great degree of numerical agreement throughout the entire range of color vision phenomena. Yet, because the experimental results which underlie the theory are fundamental, the theory is likely to provide an essentially correct description of the phenomena of color vision even if it is not quantitatively accurate. The proposed theory may therefore be thought of as a first-order approximation to a "correct" mathematical description of phenomenological color vision in much the same way as the more far-reaching and subtle Newtonian theory of dynamics is but an approximate mathematical description of the motion of material bodies.

By a "phenomenological" theory of color perception we mean a theory which interposes a transducer (whose internal structure will not be of interest to us in the present account) between an input color signal (a "light") and an output perceptual response. Moreover, this paper will be concerned only with perceptual responses which are elicited from physically steady-state color configurations (but the saccadic eye motions are of course permitted); excluded, for instance, are the results of flicker photometry experiments but the results of splitfield photometric experiments are admitted. The distinction will be of special importance in what follows because the so-called "Abney's Law" is confirmed by flicker photometric experiments but contradicted by split-field experiments[1]; the theory can be expected to agree with at most one of these results.

The perception of colors by "normal" observers differs in detail from one observer to another. To the extent that a theory of color perception is independent of the particular normal observer, it fails to be a description of the color perception of any particular real observer but instead constitutes a theory of color perception for a ficticious idealized observer. Following the thought of a previous paragraph, this remark already shows that strict quantitative agreement of the theory with experimental results of measurements of the perceptions of a small number of observers cannot be expected.

We will think of all real normal observers as approximations of varying degrees of accuracy of a ficticious ideal normal observer whose perceptual capabilities include but will not be entirely limited to the normal range of perceptual abilities exhibited by real observers. Thus, the ideal observer will be assumed capable of perceiving certain ficticious hues more highly saturated than spectral hues; perceptions of this type can be induced by selective bleaching of one of the photopigments of the retina, but the physiology of the possible production of such ficticious saturations will not be considered in adherence with our phenomenological viewpoint. Similarly, although each real human color perception transducer has a maximum admissible energy intensity beyond which the transducer experiences irreversible failure due to the destruction of its constituent parts, one may assume that the ideal human color perception transducer does not suffer this deficiency but can accept arbitrarily large input intensities without experiencing destructive failure. Finally in this regard, we are here interested only in a theory of color perception for daylight vision. The well-known effects of the quantum nature of light which become significant for inputs of low intensity will not be included in the range of applicability of the theory proposed; on the contrary, we will suppose that the ideal human color perception transducer perceives light inputs of arbitrarily low, as well as high, intensity according to the principles established by experiment for daylight vision at intermediate intensities.

3. Notation and Terminology

Our knowledge of the properties of light comes to us through the medium of one type of transducing system or another; if the system is that constituted by the human eye—brain combination, we speak of color perception, whereas if the transducer is conceived as the collection of all apparatus currently used in physical optics, we speak of the "physical" or "optical" properties of the light. The task of a theory of color perception is to relate the output signals of the ideal perception transducer to the physical properties (i.e., the output of the optical transducing system) for a common input light.

[1] Personal communication from Prof. H. G. Sperling.

We will denote lights by lower case gothic letters, $\mathfrak{x}$, $\mathfrak{y}$, etc., the corresponding physical light by the corresponding bold face font lower case roman letter, and perceptual lights by corresponding lower case roman letters. Thus

$$\mathfrak{x} \mapsto \boldsymbol{x}(\mathfrak{x}) = \boldsymbol{x}$$

denotes the correspondence of the light $\mathfrak{x}$ to the optical transducer output $\boldsymbol{x}$, and

$$\mathfrak{x} \mapsto x(\mathfrak{x}) = x$$

denotes the correspondence of $\mathfrak{x}$ to the output x of the ideal human color perception transducer. The lights $\mathfrak{x}$ are abstractions whose properties are known only in terms of the corresponding transducer outputs; they play little direct role in what follows. A theory of color perception is concerned with the relationship between $\boldsymbol{x}(\mathfrak{x})$ and $x(\mathfrak{x})$ as $\mathfrak{x}$ varies through the set of lights.

Physical optics provides a description of $\boldsymbol{x}(\mathfrak{x})$ as an element of a real Hilbert space defined on an interval $I = [\nu_{\min}, \nu_{\max}] \subset \boldsymbol{R}$. $\nu_{\min}$ and $\nu_{\max}$ are called the "minimum visible frequency" and the "maximum visible frequency" respectively. Each physical light $\boldsymbol{x}$ is an equivalence class of non-negative functions (other than the zero function) $\boldsymbol{x}: I \mapsto \boldsymbol{R}$ which differ only on a set of Lebesgue measure 0 (we will identify an equivalence class with any of its representatives without fear of confusion) such that $0 < \int_I \boldsymbol{x}^2 < \infty$ with respect to Lebesgue measure.

The value of the physical light function $\boldsymbol{x}(\mathfrak{x})(\nu) =_{def} \boldsymbol{x}(\nu)$ for $\nu \in I$ is the intensity of the physical light at frequency ν; $\int_I \boldsymbol{x}$ is the intensity of $\boldsymbol{x}$.

Denote the Hilbert space of square integrable functions on I by $\mathscr{H}$, and let $\mathscr{L} = \{\boldsymbol{x} \in \mathscr{H}: \boldsymbol{x}$ is a physical light$\}$. $\mathscr{L}$ spans $\mathscr{H}$. Addition of elements in $\mathscr{L}$ corresponds to physical superposition of the underlying abstract lights (by the use of a system of mirrors, for instance). Scalar multiplication by positive real numbers can be expressed in terms of physical superposition: this is clear for multiplication by positive integers; for multiplication by rational numbers, fractionation of a light signal by fractionation of a mirror suffices and the general case follows in the same way, or as a limit of rational fractionation.

It is known that $\mathscr{H}$ is not finite dimensional because, as Newton discovered, there exist lights — so called "spectral lights" — for which the corresponding physical light function is supported at $\nu \in I$ for any given ν (there is in fact some small spread of frequencies in the support of a spectral light, but it is not a significant restriction to take this into account), but no physical spectral light is the sum (= superposition) of distinct physical spectral lights.

Let two elements $\boldsymbol{x}, \boldsymbol{y} \in \mathscr{L}$ be called *equivalent*, denoted $\boldsymbol{x} \sim \boldsymbol{y}$, if they are perceived as the same perceptual light. This introduces a relation on $\mathscr{L}$ which is obviously reflexive and symmetric. That it is also transitive, i.e., that $\boldsymbol{x} \sim \boldsymbol{y}$ and $\boldsymbol{y} \sim \boldsymbol{z}$ implies $\boldsymbol{x} \sim \boldsymbol{z}$, is a consequence of experiment (Graham 1965, Grassmann 1853). Therefore "$\sim$" is an equivalence relation, and the equivalence classes $\mathscr{L}/\sim$ can be identified with the set $\mathscr{P}$ of perceptual lights.

The vector space addition of $\mathscr{H}$ restricted to $\mathscr{L}$, that is, superposition of lights, and multiplication of elements of $\mathscr{L}$ by positive real numbers can be transferred to the set $\mathscr{P}$ of perceptual lights by defining $\alpha\langle x\rangle+\beta\langle y\rangle=\langle\alpha x+\beta y\rangle$, where $\langle x\rangle$ denotes the perceptual equivalence class of the physical light x. These definitions can be extended to multiplication by arbitrary real numbers, from which it follows that there is a real vector space $\mathscr{V}$ spanned by $\mathscr{P}$ such that the correspondence

$$h: \mathscr{H}\rightarrow\mathscr{V}\supset\mathscr{P}$$

induced by $x\mapsto\langle x\rangle=x$ is a homomorphism of vector spaces, where equality in $\mathscr{V}$ means color matching equality. The most important single experimental fact about color perception is that $\mathscr{V}$ is finite dimensional. From this it follows that infinitely many physically distinct lights must coincide perceptually.

4. Affine Properties of Color Space

When we think of the perception of a light x we have the following observational arrangement in mind: x is presented to the observer's eye as a small central illuminated area seen against a uniformly illuminated background. The observation of x therefore implicitly assumes the presence of some background illuminant. It will further be assumed that the color perception transducer of the observer has reached a steady state before its output is sampled, by, say, a color matching procedure.

We will formulate various standard experimental results as axioms delimiting the geometry of the set $\mathscr{P}$ of perceived lights.

Axiom 1: If $x\in\mathscr{P}$ and $\alpha>0$, then $\alpha x\in\mathscr{P}$.

This axiom asserts that every positive multiple of a perceived light is a perceived light. The experimental results verify this statement only for a restricted range of positive numbers: if α is very large, the transducer will be destroyed, whereas if α is a sufficiently small positive number, quantum effects and "background noise" in the transducer conspire to destroy the relation between input and output which persists for normal daylight intensities. Axiom 1 is an idealization of the experimental results which is taken to describe the behavior of the ficticious ideal observer. Axiom 1 asserts that $\mathscr{P}$ is a cone in $\mathscr{V}$ (Newton 1704).

Axiom 2: If $x\in\mathscr{P}$, then there does not exist $y\in\mathscr{P}$ such that $x+y=0$.

That is, no superposition of perceived lights produces the absence of perceived light. Axiom 2 asserts that the cone $\mathscr{P}$ does not contain any 1 dimensional vector space.

Axiom 3: (Grassmann 1853, Helmholtz 1866) For every x, $y\in\mathscr{P}$ and $\alpha\in[0,1]$, $\alpha x+(1-\alpha)y\in\mathscr{P}$.

That is, the line segment which joins perceived lights x and y consists entirely of perceived lights. The experiment consists of superposition of lights αx and $(1-\alpha)y$ for values of $\alpha\in[0,1]$, and consequent observation that $\alpha x+(1-\alpha)y$ is always a visible light. This axiom implies (together with the previous ones) that the cone $\mathscr{P}$ is connected, hence simply connected and contractible.

Axiom 4: (Grassmann 1853) Any 4 perceived lights are linearly dependent.

That is, if $x_k \in \mathscr{P}$, $k=1, \ldots, 4$, then there are $\alpha_k \in \boldsymbol{R}$ such that

$$\sum_{k=1}^{4} \alpha_k x_k = 0.$$

It follows from Axiom 2 that not all of the x_k can be of the same sign and therefore the difference between the number of positive and negative signs is either 0 or 2. If it is 2, then one perceived light is expressed as a superposition of the remaining three; if the difference is 0, then Axiom 4 asserts that the superposition of two perceived lights matches the superposition of the remaining two. As is well known, this case does occur; not every perceived light can be expressed as a superposition of 3 linearly independent perceived lights.

Axiom 4 implies that $\dim \mathscr{V} \leq 3$. $\dim \mathscr{V}$ is a characteristic of each observer. Those observers for whom $\dim \mathscr{V} = 3$, 2, 1, and 0 are respectively called "trichromate", "dichromate", "monochromate", and "blind" observers.

Axioms 1—4 provide the affine structure of the set $\mathscr{P}$ of perceived lights, essentially following the elegant exposition of Schrödinger (1920). The axioms which follow entail metrical consequences for $\mathscr{P}$.

5. Homogeneity of Color Space

Denote by $GL(\mathscr{P})$ the group of orientation preserving linear transformations of $\mathscr{V}$ which preserve the cone $\mathscr{P}$ of perceived colors. An element $g \in GL(\mathscr{P})$ will be called a *change of background illumination.* Indeed, recalling the observational configuration envisioned in chapter 3, if $x_{\mathfrak{y}}(\mathfrak{x})$ be the perceived light (an element of $\mathscr{P}$) when $\mathfrak{x}$ is exhibited to the ideal observer displayed against the background illuminant $\mathfrak{y}$, and if $\mathfrak{y}$ is replaced by an illuminant $\mathfrak{y}'$, and $\mathfrak{x}$ by $\mathfrak{x} + \mathfrak{y}' - \mathfrak{y}$, then $x_{\mathfrak{y}}(\mathfrak{x})$ will be replaced by $x_{\mathfrak{y}'}(\mathfrak{x} + \mathfrak{y}' - \mathfrak{y})$, which is in general a perceived light distinct from $x(\mathfrak{x})$. On the perceptual level, if x_y is written for x perceived against the background y, then the map $y \overset{g}{\mapsto} y'$ induces a map $x_y \overset{g}{\mapsto} (x + y' - y)_{y'}$ and evidently $g(\mathscr{P}) \subset \mathscr{P}$, that is, the perceived light remains a perceived light. Conversely, if the illuminant y replaces y' in this way, the corresponding induced map is g^{-1} and $g^{-1}(\mathscr{P}) \subset \mathscr{P}$, whence g preserves $\mathscr{P}$. Finally, each y' can be expressed as the image $g(y)$ for some endomorphism of $\mathscr{V}$, so $g \in GL(\mathscr{P})$.

Examples of changes of background illumination are provided by the varying illumination provided by sunlight as a function of the time of day and the cloud cover, by the transformation induced by replacement of a daylight illuminant by a typical incandescent or fluorescent illuminant, and by the transformation of background induced by the use of sunglasses or other filtering media.

Axiom 5: $\mathscr{P}$ is locally homogeneous with respect to changes of background illumination.

Consider $\mathscr{V}$ with its usual topology. Axiom 5 asserts that every $x \in \mathscr{P}$ has an open neighborhood $U \subset \mathscr{P}$ such that every $y \in U$ can be expressed as $y = gx$ for some $g \in GL(\mathscr{P})$. Thus Axiom 5 implies in particular that $\mathscr{P}$ is open in $\mathscr{V}$ and therefore inherits the structure of a differentiable manifold from $\mathscr{V}$. The observational interpretation of this axiom is simply that any perceived light x can be transformed into any "sufficiently near" light y by an appropriate change of background illumination, and is known to be true, at least for neighborhoods of points in $\mathscr{P}$ which lie near the image $x(c)$ of any typical daylight illuminant c. Compare the discussion of homogeneity in section 2.

If x and y are perceived lights and L is the line segment which joins x to y, then one might expect to be able to pass from x to y by a change of background illumination g constructed by selecting a sequence of perceived lights $x = x_1$, $x_2, x_3, \ldots, x_n, x_{n+1} = y$, $x_k \in L$, such that consecutive perceived lights x_k, x_{k+1} are so close to each other that there exists a change of background illumination g_k which carries x_k onto x_{k+1}. If such a sequence exists, then the composition

$$g = g_n \circ g_{n-1} \circ \cdots \circ g_1$$

will be a change of background illumination which carries x to y, so $\mathscr{P}$ would in fact be (globally) homogeneous with respect to $GL(\mathscr{P})$. The line segment L which joins x to y in $\mathscr{P}$ is compact. Let U_z be a locally homogeneous neighborhood of $z \in L$, whose existence is assured by Axiom 5. The sets $\{U_z\}$ cover L and hence by compactness there are finitely many points $x_k \in L$, $1 \leq k \leq n+1$ such that $L \subset \bigcup_{k=1}^{n+1} U_{x_k}$ and the x_k are a sequence of the desired form. Hence $\mathscr{P}$ is (globally) homogeneous if and only if it is locally homogeneous. We therefore formulate

Axiom 5′: $\mathscr{P}$ is (globally) homogeneous with respect to changes of background illumination.

$GL(\mathscr{P})$ is a subgroup of $GL(\mathscr{V})$ and is therefore a Lie group. Hence, by standard results in the theory of homogeneous spaces, $\mathscr{P}$ can be identified with the homogeneous space $GL(\mathscr{P})/K$, where K is isomorphic to the subgroup of $GL(\mathscr{P})$ which leaves some point of $\mathscr{P}$ fixed, hence to a closed subgroup of the orthogonal group, and consequently to a compact subgroup of $GL(\mathscr{P})$. Moreover, since the map $x \mapsto \alpha x$, $\alpha \in \boldsymbol{R}^+$ preserves $\mathscr{P}$, it follows that each $g \in GL(\mathscr{P})$ can be uniquely expressed in the form $g = \alpha \circ h$ where $\alpha \in \boldsymbol{R}^+$ and $h \in SL(\mathscr{P}) = GL(\mathscr{P}) \cap SL(\mathscr{V})$; $SL(\mathscr{V})$ denotes the elements of $GL(\mathscr{V})$ whose matrix representative relative to a basis for $\mathscr{V}$ has determinant $+1$. It follows that $GL(\mathscr{P}) = \boldsymbol{R}^+ \times SL(\mathscr{P})$ and that K can be identified with a compact subgroup of $SL(\mathscr{P})$.

Now the dimension axiom, Axiom 4, insures that $SL(\mathscr{P})$ is isomorphic to a subgroup of $SL(1, \boldsymbol{R})$, $SL(2, \boldsymbol{R})$, or $SL(3, \boldsymbol{R})$ according as the dimension of $\mathscr{V}$ is 1, 2, or 3. As usual, $SL(n, R)$ denotes the group of $n \times n$ real matrices with determinant equal to $+1$. Unless otherwise stated, let us restrict our considerations to trichromate observers, i.e., to the case $\dim \mathscr{V} = 3$. From Helgason (1962), $\dim SL(3, \boldsymbol{R}) = 8$, and

$$3 = \dim \mathscr{P} = \dim \boldsymbol{R}^+ \times SL(\mathscr{P})/K = 1 + \dim SL(\mathscr{P}) - \dim K; \tag{1}$$

hence

$$8 \geq \dim SL(\mathscr{P}) = 2 + \dim K \tag{2}$$

where K is a compact subgroup of $SL(\mathscr{P}) \subset SL(3, \boldsymbol{R})$.

The simple real Lie groups have been classified. This permits one to determine, up to isomorphism, the possible forms which $SL(\mathscr{P})$ and K may have. First remark that the contractibility of $\mathscr{P}$ implies that $SL(\mathscr{P})$ is the exponential of its Lie algebra and the classification of Lie algebras (Jacobson 1962, Helgason 1962) shows that the simple noncompact Lie groups of dimension ≤ 8 are just $SL(2, \boldsymbol{R})$ and $SL(3, \boldsymbol{R})$. Further, a nilpotent Lie algebra can be considered as a Lie algebra of upper triangular matrices, so the corresponding irreducible unipotent Lie groups are of the form

$$T_n = \left\{ \begin{pmatrix} 1 & & \alpha_{ij} \\ & \ddots & \\ 0 & & 1 \end{pmatrix} : \alpha_{ij} \in \boldsymbol{R},\ 1 \leq i < j \leq n \right\}.$$

Since $\dim T_n = \dfrac{n(n-1)}{2}$, only the range $2 \leq n \leq 4$ is relevant. T_n has no compact subgroups. Hence there is a semi-simple group S and integers n_i such that

$$\dim SL(\mathscr{P}) = \dim S + \dim (T_{n_1} \times \ldots \times T_{n_k}),\quad S \text{ semi-simple or absent,} \tag{3}$$

K is a compact subgroup of the semi-simple group S, and the dimensionalities of these groups must satisfy

$$8 \geq \dim S + \sum_{i=1}^{k} \frac{n_i(n_i-1)}{2} = 2 + \dim K. \tag{4}$$

Then $S = \emptyset$, $S = SL(2, \boldsymbol{R})$, $S = SL(3, \boldsymbol{R})$, or $S = SL(2, \boldsymbol{R}) \times SL(2, \boldsymbol{R})$.

If $S = \emptyset$, then $K = \emptyset$ and (4) reduces to

$$\sum_{i=1}^{k} \frac{n_i(n_i-1)}{2} = 2$$

with the unique solution $n_1 = n_2 = 2$, $k = 2$. Hence $SL(\mathscr{P})/K \simeq T_2 \times T_2 \simeq \boldsymbol{R}^+ \times \boldsymbol{R}^+$, so as a homogeneous space,

$$\mathscr{P} = GL(\mathscr{P})/K \simeq \boldsymbol{R}^+ \times SL(\mathscr{P})/K \simeq \boldsymbol{R}^+ \times \boldsymbol{R}^+ \times \boldsymbol{R}^+. \tag{5}$$

We will show below that this case contains Helmholtz's original model for $\mathscr{P}$, as well as Stiles' modification of Helmholtz's model (Stiles 1946; Stiles and Wyszecki 1967).

Next suppose $S = SL(2, \boldsymbol{R})$. Then $\dim S = 3$ and (4) becomes

$$5 \geq \sum_{i=1}^{k} \frac{n_i(n_i-1)}{2} = \dim K - 1.$$

The compact subgroups of $SL(2, \boldsymbol{R})$ are isomorphic to the group with one element (hence of dimension 0) or to the orthogonal group $SO(2)$ of matrices of the form

$$\begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}, \quad 0 \le \theta < 2\pi,$$

of dimension 1. In the first case,

$$5 \ge \sum_{i=1}^{k} \frac{n_i(n_i-1)}{2} = -1,$$

a contradiction. In the second case, $\dim K = 1$ and

$$5 \ge \sum_{i=1}^{k} \frac{n_i(n_i-1)}{2} = 0$$

is satisfied by choosing all $n_i = 0$, that is, by the realization

$$\mathscr{P} = GL(\mathscr{P})/K \simeq \boldsymbol{R}^+ \times SL(\mathscr{P})/K \simeq \boldsymbol{R}^+ \times SL(2, \boldsymbol{R})/SO(2). \tag{6}$$

Thus, in this case, $\mathscr{P}$ is equivalent as a homogeneous space to the product of $\boldsymbol{R}^+$ (or, what is the same as a homogeneous space, $\boldsymbol{R}$) and the Poincaré-Lobachevsky two dimensional space of constant negative curvature. The implications of this model for color perception will be explored in a later section.

In order to complete the classification of realizations of $\mathscr{P}$ as a homogeneous space of Lie groups, consider $S = SL(3, \boldsymbol{R})$. Then (4) becomes

$$8 \ge 8 + \sum \frac{n_i(n_i-1)}{2} = 2 + \dim K;$$

hence, the n_i must all be 0 and K must be a compact subgroup of $SL(3, \boldsymbol{R})$ of dimension 6. But a maximal compact subgroup of $SL(3, \boldsymbol{R})$ is $SO(3)$ with $\dim SO(3) = 3$, which contradicts the existence of K. The final case is disposed of similarly.

These results may be summarized by the assertion that Axioms 1—5 imply $\mathscr{P}$ is a homogeneous space equivalent either to $\boldsymbol{R}^+ \times \boldsymbol{R}^+ \times \boldsymbol{R}^+$, or to $\boldsymbol{R}^+ \times SL(2, \boldsymbol{R})/SO(2)$.

6. Invariant Color Metrics

We wish to introduce a Riemannian metric on $\mathscr{P}$ which will provide a measure of the (perceptual) dissimilarity of perceived lights. It is naturally desirable that the metric imposed on $\mathscr{P}$ be compatible with Axioms 1 – 5, that is, with the structure of $\mathscr{P}$ as a homogeneous space. In order to motivate the choice, consider a pair of perceived lights x and y viewed with respect to a background illuminant b. It is common experience that a change of background illumination $b \mapsto b'$ produces a change in the perceived lights, $x \mapsto x'$, $y \mapsto y'$, but that, at least if the perceptual distance between b and b' is not great, the perceptual distance between x' and y' is the same as the perceptual distance between x and y; recall the discussion of this point in chapter 2, and the examples adduced there.

Let the Riemannian metric which provides the measure of perceptual dissimilarity be given in the usual infinitesimal form by a quadratic form ds^2. Following Schrödinger (1920), suppose that the measure of dissimilarity of x and $y \in \mathscr{P}$ is

$$d(x, y) = \int ds \tag{7}$$

where the integral is calculated along a geodesic arc which connects x and y. In order that this definition be consistent, it is necessary that all geodesic arcs connecting x and y be of the same length, a condition which will be insured in what follows by the uniqueness of geodesic arcs between point pairs in $\mathscr{P}$. If the relative perceptual distances between point pairs remains unchanged upon changes of background illumination, as the previous paragraph suggests, then

$$d(gx, gy) = d(x, y) \quad \text{for} \quad x, y \in \mathscr{P},\ g \in GL(\mathscr{P});$$

the distance function is a point pair invariant with respect to the group $GL(\mathscr{P})$ which acts transitively on $\mathscr{P}$. We formalize this as

Axiom 6: The Riemannian metric on $\mathscr{P}$ which measures perceptual dissimilarity is a $GL(\mathscr{P})$ — invariant metric.

This axiom determines the perceptual metric (as we shall refer to the metric described in Axiom 6). Indeed, if $x, y \in \mathscr{P}$ and $g \in GL(\mathscr{P})$ carries $x \mapsto gx = y$, and if G_x is the metric on the tangent space $\mathscr{T}_x$ to $\mathscr{P}$ at x induced by the given Riemannian metric on $\mathscr{P}$, then the differential dg of g induces an isomorphism of $\mathscr{T}_x$ on $\mathscr{T}_{gx}$, and G_{gx} is defined by

$$G_{gx}(dgX) = G_x(X) \quad \text{for} \quad X \in \mathscr{T}_x.$$

Since $GL(\mathscr{P})$ is transitive on $\mathscr{P}$, the metric is determined everywhere by the metric G_x on $\mathscr{T}_x$ for a fixed but arbitrary $x \in \mathscr{P}$. Let us identify x with the coset K in the realization of $\mathscr{P}$ as the homogeneous space $GL(\mathscr{P})/K$. Then evidently $gx = x$ if $g \in K$ and therefore the metric G_x on the tangent space $\mathscr{T}_x$ must be K-invariant, that is, $G_x(dgX) = G_x(X)$ for $g \in K$ and $X \in \mathscr{T}_x$. There are two cases to consider, corresponding to the two representations for $\mathscr{P}$ as a homogeneous space. First, if $\mathscr{P} = \boldsymbol{R}^+ \times SL(2, \boldsymbol{R})/SO(2)$, then $\mathscr{T}_x = \boldsymbol{R} \oplus \mathscr{T}_x'$ where $\mathscr{T}_x'$ is the 2-dimensional subspace of $\mathscr{T}_x$ which is tangent to $SL(2, \boldsymbol{R})/SO(2)$ at $K = SO(2)$. The restriction of the metric to $\mathscr{T}_x'$ must be $SO(2)$-invariant, that is, invariant with respect to rotation about the origin in the 2-dimensional vector space $\mathscr{T}_x'$, and consequently it is a multiple of the Euclidean metric. It follows that G_x is the sum of a 1-dimensional and a 2-dimensional Euclidean metric and therefore the perceptual metric on $\mathscr{P}$ is unique up to selection of the units of measure on each of the factors of $\mathscr{P} = \boldsymbol{R}^+ \times SL(2, \boldsymbol{R})/SO(2)$. In the next section we will obtain an explicit description of a conveniently normalized form of this metric.

In the second case, $\mathscr{P} = \boldsymbol{R}^+ \times \boldsymbol{R}^+ \times \boldsymbol{R}^+$, so $K = \emptyset$ and K-invariance does not provide any restrictions on the metric. However, a $G(\mathscr{P})$-invariant metric must be the sum of metrics on each factor which are $\boldsymbol{R}^+$-invariant; since an $\boldsymbol{R}^+$-invariant metric on $\boldsymbol{R}^+$ is determined by a positive constant on the tangent space at one point, it is clear that all $\boldsymbol{R}^+$-invariant metrics on $\boldsymbol{R}^+$ are proportional. But

$ds^2 = \left(\frac{dx}{x}\right)^2$ is patently an R^+-invariant metric on R^+. Hence, the general $GL(\mathscr{P})$-invariant metric on $\mathscr{P} = R^+ \times R^+ \times R^+$ is

$$ds^2 = \alpha_1 \left(\frac{dx_1}{x_1}\right)^2 + \alpha_2 \left(\frac{dx_2}{x_2}\right)^2 + \alpha_3 \left(\frac{dx_3}{x_3}\right)^2, \tag{8}$$

where the α_k are positive constants. This is precisely Stiles' generalization of Helmholtz' metric (Helmholtz 1892, Stiles 1946); the Helmholtz metric corresponds to $\alpha_1 = \alpha_2 = \alpha_3 = 1$. Note that the metric (8) admits unique geodesic arcs connecting arbitrary point pairs in $\mathscr{P}$, which insures that the definition (7) of perceptual dissimilarity of distinct perceived lights is consistent.

7. Realization of $R^+ \times SL(2,R)/SO(2)$

A convenient realization of $R^+ \times SL(2, R)/SO(2)$ can be given as follows. Let $\mathscr{P}$ denote the set of all 2×2 symmetric real matrices x which are positive definite. Denote the determinant of x by $|x|$. Let $\mathscr{N}$ denote the subset of $\mathscr{P}$ which consists of matrices of determinant 1. Then the decomposition $x = |x| \left(\frac{x}{|x|}\right)$ shows that $\mathscr{P} = R^+ \times \mathscr{N}$. It is well-known that $\mathscr{N}$ is isomorphic to $SL(2, R)/SO(2)$. In fact, $R^+ \times SL(2, R) = GL(2, R)$ acts on $\mathscr{P}$ as follows: if $A \in GL(2, R)$ and A^t denotes the transposed matrix, then $\mathscr{P} \ni x \mapsto AxA^t$. Denote the trace of a matrix x by $tr\, x$. Then

$$ds^2 = \mathrm{tr}\,(x^{-1}\, dx\, x^{-1}\, dx) \tag{9}$$

defines a conveniently normalized form of the perceptual metric on $\mathscr{P}$; the general form of the metric will be given in section 8. Indeed, under the map

$$x \to AxA^t = gx, \text{ find } x^{-1} = (A^t)^{-1}\, x^{-1}\, A^{-1},\ dgx = AdxA^t,$$

whence

$$\mathrm{tr}((gx)^{-1}\, d(gx)\,(gx)^{-1}\, d(gx)) = \mathrm{tr}\,((A^t)^{-1}\, x^{-1}\, dxx^{-1}\, dxA^t) = \mathrm{tr}\,(x^{-1}\, dxx^{-1}\, dx).$$

The axioms thus far presented have determined two distinct classes of homogeneous spaces which are candidates for the space of perceived lights, and corresponding perceptual metrics which are uniquely determined up to the choice of the unit, or scale, of measurement. One of these is Stiles' generalization of Helmholtz' model.

8. Brightness

The next task is to identify perceived *brightness* of lights in terms of the geometrical structures of the two types of Riemannian manifolds which satisfy Axioms 1—6, and to compare the consequences of that identification with the conclusions of psychophysical studies.

If x is a physical light and $\alpha>0$, then the physical light αx differs from x only by its intensity. The corresponding perceived distinction is called *brightness*; the measure of relative brightness of x and αx should, according to the definition of the perceptual metric, be measured by the perceptual distance between x and αx in $\mathscr{P}$. That is, the relative brightness is (up to the choice of a unit of measurement)

$$d(\alpha x, x) = \int_x^{\alpha x} ds.$$

For the metric (8) we find, since $d(\alpha x) = x d\alpha$ on the geodesic which connects x to αx,

$$d(\alpha x, x) = \sqrt{\alpha_1^2+\alpha_2^2+\alpha_3^2} \int_1^{\alpha} \frac{d\tau}{\tau},$$

while for the second model type, with metric given by (9)

$$d(\alpha x, x) = \sqrt{\operatorname{tr}\begin{pmatrix}1&0\\0&1\end{pmatrix}} \int_1^{\alpha} \frac{d\tau}{\tau} = \sqrt{2} \int_1^{\alpha} \frac{d\tau}{\tau}$$

Hence, in either case[2],

$$d(\alpha x, x) = \varkappa \log \alpha, \quad \varkappa = \begin{cases} \sqrt{\alpha_1^2+\alpha_2^2+\alpha_3^2} & \text{for metric (8);} \\ \sqrt{2} & \text{for metric (9).} \end{cases} \tag{10}$$

Equation (10) is the statement of the psychophysical "law" of Fechner (Fechner 1877; cp. Ekman 1959, Hecht 1935, Stevens 1957, 1970) applied to the brightness sensation, and is a consequence of the axioms given above for physical lights which differ solely in intensity.

We are now prepared to consider the difficult problem of heterochromatic brightness comparison. Let $x, y \in \mathscr{P}$. Since x and y do not in general lie on the same ray emanating from $0 \in \mathscr{V}$, that is, since x and y differ, in general, by more than a multiplicative intensity factor, the perceptual distance $d(x, y)$ will measure chromatic as well as brightness distinctions. If $b(x, y)$ denotes the measure of brightness distinction, that is, the brightness of x relative to y, then in general $b^2(x, y) \leq d^2(x, y)$ since brightness accounts for only a portion of the perceived distinction.

The realizations of $\mathscr{P}$ obtained in section 5 show that in all cases, $\mathscr{P} \cong \boldsymbol{R}^+ \times \mathscr{M}$, where $\mathscr{M} \cong \boldsymbol{R}^+ \times \boldsymbol{R}^+$ or $\mathscr{M} \cong SL(2, \boldsymbol{R})/SO(2, \boldsymbol{R})$, and "$\cong$" means "isometric to". It follows that any $x \in \mathscr{P}$ has a unique representation as

$$x = (\xi, u), \quad \xi \in \boldsymbol{R}^+ \text{ and } u \in \mathscr{M}. \tag{11}$$

For $\mathscr{M} = \boldsymbol{R}^+ \times \boldsymbol{R}^+$ this representation is given explicitly by

[2] More precisely, eq. (10) exhibits the *signed distance* from x to αx, which is positive if $\alpha>1$ but negative if $\alpha<1$, corresponding to the notions "brighter than" and "less bright than" The distance itself is given by the absolute value of the right side of (10).

$$\xi = (x_1^{\alpha_1} x_2^{\alpha_2} x_3^{\alpha_3})^{1/3}$$
$$x_1^\sigma = \xi u_1 \qquad \text{with } \sigma = \frac{\alpha_1 + \alpha_2 + \alpha_3}{3} \tag{12}$$
$$x_2^\sigma = \xi u_2$$

In the (ξ, u) coordinates, the perceptual metric (8) becomes

$$\sigma^2 \, ds^2 = 3\,\sigma \left(\frac{d\xi}{\xi}\right)^2 + \tag{13}$$
$$+ \left[\left(\alpha_1 + \frac{\alpha_1^2}{\alpha_3}\right)\left(\frac{du_1}{u_1}\right)^2 + \frac{2\,\alpha_1\,\alpha_2}{\alpha_3}\left(\frac{du_1}{u_1}\right)\left(\frac{du_2}{u_2}\right) + \left(\alpha_2 + \frac{\alpha_2^2}{\alpha_3}\right)\left(\frac{du_2}{u_2}\right)^2 \right],$$

which shows that the perception of the ξ-variable is measured independently of perception of the variable $u \in \mathcal{M}$. If x is fixed and $\alpha > 0$, then as we already know, x and αx differ only in brightness. But in the coordinate system (12), if $x = (\xi, u)$, it follows that $\alpha x = (\alpha^\sigma \xi, u)$ whence the brightness difference is expressed entirely in terms of the ξ-coordinate in the coordinate system (12).

Similarly, if $\mathcal{M} \cong SL(2, R)/SO(2)$ and if $\mathcal{P}$ is represented as the set of positive definite 2×2 real symmetric matrices $x = \begin{pmatrix} x_1 & x_3 \\ x_3 & x_2 \end{pmatrix}$, then an explicit representation which exhibits the product structure of $\mathcal{P}$ is provided by the new coordinatization $x = (\xi, u)$ where $u \in \mathcal{M}$ and

$$\xi = |x|^{1/2},$$
$$u_1 = x_3/x_2, \tag{15}$$
$$u_2 = |x|^{1/2}/x_2 > 0.$$

This corresponds to the decomposition of the matrix x as

$$x = \xi \begin{pmatrix} (u_1^2 + u_2^2)/u_2 & u_1/u_2 \\ u_1/u_2 & 1/u_2 \end{pmatrix}. \tag{16}$$

In these coordinates, the normalized perceptual metric (9) assumes the form

$$ds^2 = \left(\frac{d\xi}{\xi}\right)^2 + \left[\frac{(du_1)^2 + (du_2)^2}{(u_2)^2}\right]; \tag{17}$$

the general metric is evidently given by

$$ds^2 = \alpha_1 \left(\frac{d\xi}{\xi}\right)^2 + \alpha_2 \left[\frac{(du_1)^2 + (du_2)^2}{(u_2)^2}\right]$$

where α_1, α_2 are positive constants, which shows that in this case also, perception of the ξ-variable is independent of perception of the variables $u \in \mathcal{M}$. Moreover, if x is fixed and $\alpha > 0$, then x and αx differ only in brightness and have coordinates

$$x = (\xi, u), \quad \alpha x = (\alpha\xi, u)$$

in the coordinate system (15), which shows that the brightness difference is expressed solely in terms of the ξ-variable.

These observations motivate the following definition of relative brightness of heterochromatic perceived lights. If $x, y \in \mathscr{P}$, then the *brightness of x relative to y* is

$$b(x, y) = \varkappa \int_{\eta}^{\xi} \frac{d\tau}{\tau} = \varkappa \log \frac{\xi}{\eta}. \tag{18}$$

where $x = (\xi, u)$ and $y = (\eta, v)$ are the coordinates of x and y in the appropriate coordinate system (12) or (15), and $\varkappa$ is a constant chosen so that (18) agrees with (10) when x and y are proportional.

For our later purposes it will be useful to remark that if one writes $z = u_1 + iu_2$, $\bar{z} = u_1 - iu_2$, then the restriction of the metric (17) to $\mathscr{M}$, i.e., the expression in square brackets on the right side of (17), becomes $\frac{dz d\bar{z}}{(\operatorname{Im} z)^2}$ where $\operatorname{Im} z$ denotes the imaginary part of z. Since x is positive definite in (15), it follows that $\operatorname{Im} z = u_2 > 0$, so the metric $\frac{dz d\bar{z}}{(\operatorname{Im} z)^2}$ equips the "upper half plane" $\{z: \operatorname{Im} z > 0\}$ with the usual metric with respect to which it is isometric to the Poincaré model of the (non-compact) two dimensional space of constant negative curvature.

9. Surfaces of Constant Brightness

The famous "law" of Abney (Abney and Festing 1886, Abney 1913, Graham 1965) is the assertion that perceived brightness is a linear function of the superposition of physical colors; it can be analytically formulated as follows.

Let $c \in \mathscr{P}$ denote a fixed reference perceived light and let $b(x, c)$ denote, as above, the brightness of x relative to c. Then Abney's Law states that for any $x, y \in \mathscr{P}$ and $\alpha, \beta \in \boldsymbol{R}^+$,

$$b(\alpha x + \beta y, c) = \alpha b(x, c) + \beta b(y, c). \tag{19}$$

Under the hypothesis that Abney's Law (19) is valid, the set of perceived lights with a given brightness is the intersection of a plane in $\mathscr{V}$ with the cone of perceived lights $\mathscr{P}$. Indeed, if $e_1, e_2, e_3 \in \mathscr{P}$ is a basis for $\mathscr{V}$, then any $x \in \mathscr{P}$ can be expressed in the form

$$x = \sum_{k=1}^{3} \xi_k e_k$$

so, for lights of brightness β,

$$b(x, c) = \sum \xi_k b(e_k) = \beta, \quad x \in \mathscr{P}, \tag{20}$$

the equation of a plane.

It is now known that Abney's Law is not valid for color matching experiments, although it does appear to agree with flicker photometric measurements (Graham 1965, Sperling, Stiles and Wyszecki 1967).

First consider the perceptual metric (8) and its decomposition (12). Let the coordinates of the reference standard $c \in \mathscr{P}$ be $c = (\gamma, v)$, and let $x = (\xi, u)$ in this representation. Then the brightness of x relative to c is, according to (18),

$$b(x, c) = \varkappa \log \frac{\xi}{\gamma} \tag{21}$$

where $\varkappa$ is a constant. The surface of fixed brightness β is the set of $x \in \mathscr{P}$ which satisfy $b(x, c) = \beta$; from (12), this is

$$\varkappa \log \left(\frac{x_1^{\alpha_1} x_2^{\alpha_2} x_3^{\alpha_3}}{\gamma} \right)^{1/3} = \beta$$

with solution

$$x_1^{\alpha_1} x_2^{\alpha_2} x_3^{\alpha_3} = \gamma e^{3\beta/\varkappa} = \text{const.}; \tag{22}$$

these are precisely the surfaces of constant brightness found by Stiles; they are isometric to $\mathscr{M} \simeq \boldsymbol{R}^+ \times \boldsymbol{R}^+$.

For the perceptual metric (9), the use of (15) in (21) yields

$$\varkappa \log \frac{|x|^{1/2}}{\gamma} = \beta$$

whence the surfaces of constant brightness are the hyperboloids

$$|x| = \gamma e^{2\beta/\varkappa} = \text{const.}, \tag{23}$$

which are isometric to $\mathscr{M}$, that is to the Poincaré-Lobachevsky space of constant negative curvature.

The first case, eq. (22), has been fully treated (Stiles and Wyszecki 1957) so we will turn our attention to (23). Select the reference standard c and the unit of measure of brightness so that $\gamma = \frac{\varkappa}{2} = 1$. Then from (23),

$$\beta = \beta(x) = \log |x|.$$

Consequently

$$e^{\beta(x+y)} = |x + y| = |x^{1/2} (\mathbb{1} + x^{-1/2} y x^{-1/2}) x^{1/2}|$$

(where $\mathbb{1}$ denotes the unit matrix and $x^{1/2}$ is the unique (matrix) square root of $x \in \mathscr{P}$ which lies in $\mathscr{P}$.)

$$= |x| \; |\mathbb{1} + u|$$

with $u = x^{-1/2} y x^{-1/2} \in \mathscr{P}$. u is positive definite, hence has positive eigenvalues λ_1, λ_2, so $|\mathbb{1} + u| = (1 + \lambda_1)(1 + \lambda_2) = 1 + (\lambda_1 + \lambda_2) + \lambda_1 \lambda_2 = 1 + \operatorname{tr} u + |u|$ and

$$e^{\beta(x+y)} = |x| \, (1 + \operatorname{tr} u + |u|)$$

$$=|x|\left(1+\operatorname{tr} x^{-1} y+\frac{|y|}{|x|}\right)$$

$$=|x|\left(\frac{1}{|y|}+\frac{1}{|y|}\operatorname{tr} x^{-1} y+\frac{1}{|x|}\right)|y|$$

$$=e^{\beta(x)}\left(\frac{1}{|y|}+\frac{1}{|y|}\operatorname{tr} x^{-1} y+\frac{1}{|x|}\right)e^{\beta(y)}.$$

Therefore,

$$\beta(x+y)=\beta(x)+\beta(y)+\log\left(\frac{1}{|x|}+\frac{1}{|y|}+\frac{\operatorname{tr} x^{-1} y}{|y|}\right). \tag{24}$$

Now let $a\in\mathscr{P}$ be a fixed perceived light and let

$$g: x\mapsto a^{1/2} x a^{1/2},\ g\in GL(\mathscr{P}) \tag{25}$$

be a change of background illumination. Then $|a^{1/2} xa^{1/2}|=|a|\ |x|$ so

$$\beta(\alpha^{1/2} xa^{1/2})=\log|a^{1/2} xa^{1/2}|=\log|a|+\log|x|,\ \text{i. e.,}$$

$$\beta(a^{1/2} xa^{1/2})=\beta(a)+\beta(x). \tag{26}$$

Hence, if $x\mapsto a^{1/2} xa^{1/2}$ is a change of background illumination if $x\in\mathscr{P}$, then the brightness of x (relative to the standard c) is changed by an additive constant which depends only on the change of background illumination, but not on the light $x\in\mathscr{P}$. This result assumes added significance because it can be proved that every mapping in $GL(\mathscr{P})=\boldsymbol{R}^+\times SL(2,\boldsymbol{R})/SO(2)$ is a composition of transformations of the type (25), whence the conclusion drawn is valid for all changes of background illumination. This is, we believe, the correct context for Abney's Law. A similar result holds for the first model, for which $GL(\mathscr{P})\simeq\boldsymbol{R}^+\times\boldsymbol{R}^+\times\boldsymbol{R}^+$.

10. Jordan Algebras and Brightness

In the preceding sections, most arguments were carried through twice — once for each class of model of the space $\mathscr{P}$ of perceived colors. It may be worthwhile to indicate how both models can be treated in a uniform way.

If $\mathfrak{A}$ is any finite dimensional real vector space, then $\mathfrak{A}$ is said to be a *Jordan algebra* if there is a bilinear composition $\mathfrak{A}\times\mathfrak{A}\to\mathfrak{A}$, $(a,b)\mapsto ab$, such that $ab=ba$ and $a(a^2 b)=a^2(ab)$ for all $a,b\in\mathfrak{A}$. $\boldsymbol{R}$ is a Jordan algebra with respect to its usual structure as a field, while the vector space $\mathfrak{H}(r,\boldsymbol{R})$ of $r\times r$ real symmetric matrices is turned into a Jordan algebra by the introduction of the bilinear composition $(x,y)\mapsto(x\cdot y+y\cdot x)/2$, where "$\cdot$" denotes matrix multiplication; $\boldsymbol{R}=\mathfrak{H}(1,\boldsymbol{R})$. The $\mathfrak{H}(r,\boldsymbol{R})$ are instances of the class of so-called "formally real" or "compact real" Jordan algebras (Braun and Koecher 1966, Koecher 1962).

If $\mathfrak{A}$ is a Jordan algebra, then the bilinear composition induces a mapping $L: \mathfrak{A} \to \mathrm{Hom}(\mathfrak{A}, \mathfrak{A})$ defined by

$$L(a)\,b = ab \quad \text{for } a, b \in \mathfrak{A}. \tag{27}$$

The endomorphism

$$P(a) = 2\,L^2(a) - L(a^2), \quad a \in \mathfrak{A} \tag{28}$$

is called the "quadratic representation". If $\mathfrak{A}$ is a simple Jordan algebra of dimension d and rank r (Braun and Koecher 1966), then

$$\sigma(a) = \frac{r}{d}\ \mathrm{tr}\, L(a), \quad a \in \mathfrak{A} \tag{29}$$

is called the *reduced trace* of a, and

$$|a| = \{\det P(a)\}^{r/d}, \quad a \in \mathfrak{A} \tag{30}$$

is called the *reduced norm* of a. If $\mathfrak{A} = \boldsymbol{R}$, then

$$\sigma(a) = |a| = a, \quad a \in \mathfrak{A}; \tag{31}$$

if $\mathfrak{A} = \mathfrak{H}(r, \boldsymbol{R})$, then

$$\sigma(a) = \mathrm{tr}\, a,\ |a| = \text{determinant of } a, \quad a \in \mathfrak{A} \tag{32}$$

Hereafter we consider only compact real Jordan algebras $\mathfrak{A}$. $\mathfrak{A}$ has a unit element, which will be denoted by c; for $\mathfrak{A} = \mathfrak{H}(r, \boldsymbol{R})$, $c = \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix}$ is the unit matrix. The exponential function associated with $\mathfrak{A}$ is defined by

$$\exp a = \sum_{n=0}^{\infty} a^n/n! \qquad a^0 = c, \ \ a \in \mathfrak{A}. \tag{33}$$

Set $\exp \mathfrak{A} = \{\exp a : a \in \mathfrak{A}\}$; then $\exp \mathfrak{A}$ is a strictly convex open cone in $\mathfrak{A}$. If $\mathfrak{B} = \mathfrak{A}_1 \oplus \mathfrak{A}_2$ is a direct sum, then $\exp b = (\exp a_1)(\exp a_2)$ where $b = a_1 \oplus a_2 \in \mathfrak{B}$.

Observe that for $\alpha > 0$, the map $a \mapsto a/\alpha$ is an isomorphism of $\boldsymbol{R}$ onto a Jordan algebra $\mathfrak{A}_\alpha$ with unit element $1/\alpha$, and that

$$\exp \mathfrak{A}_\alpha = \{\exp \alpha a : a \in \mathfrak{A}\} = \{(\exp a)^\alpha : a \in \mathfrak{A}\} = \{x^\alpha : x \in \exp \mathfrak{A}\}.$$

Writing $\mathfrak{A}_{\alpha_1\alpha_2\alpha_3} = \mathfrak{A}_{\alpha_1} \oplus \mathfrak{A}_{\alpha_2} \oplus \mathfrak{A}_{\alpha_3}$, it follows that

$$\begin{aligned} \exp \mathfrak{A}_{\alpha_1\alpha_2\alpha_3} &= \{(x_1^{\alpha_1}, x_2^{\alpha_2}, x_3^{\alpha_3}) : x_i \in \boldsymbol{R}^+\}. \\ &= \boldsymbol{R}^+ \times \boldsymbol{R}^+ \times \boldsymbol{R}^+, \end{aligned} \tag{34}$$

and

$$\exp \mathfrak{H}(2, \boldsymbol{R}) = \left\{ \begin{pmatrix} x_1 & x_3 \\ x_3 & x_2 \end{pmatrix} = x : x \text{ is positive definite} \right\}; \tag{35}$$

thus

$$\exp \mathfrak{A} = \mathscr{P} = \text{space of perceived colors} \tag{36}$$

$$\text{if } \mathfrak{A} = \mathfrak{A}_{\alpha_1\alpha_2\alpha_3} \text{ or } \mathfrak{A} = \mathfrak{H}(2, \boldsymbol{R}).$$

The group $GL(\exp \mathfrak{A})$ is generated by the maps $P(a)$ for $a \in \mathfrak{A}$; $\exp \mathfrak{A}$ is a homogeneous space of $GL(\exp \mathfrak{A})$ and an $GL(\exp \mathfrak{A})$-invariant metric on $\exp \mathfrak{A}$ is given by

$$ds^2 = \sigma\left((P^{-1}(x)\,dx)\,dx\right). \tag{37}$$

This metric coincides with (8) if $\mathfrak{A} = \mathfrak{A}_{\alpha_1\alpha_2\alpha_3}$, and it coincides with (9) if $\mathfrak{A} = \mathfrak{H}(2, \boldsymbol{R})$.

From the definition (27), it follows that for $\mathfrak{A} = \mathfrak{H}(r, \boldsymbol{R})$,

$$P(a)\,b = a \cdot b \cdot a \quad \text{for } a, b \in \mathfrak{H}(r, \boldsymbol{R}), \tag{38}$$

with "$\cdot$" denoting matrix multiplication as before.

With the unification provided by the concept of a Jordan algebra, the arguments concerning brightness which were presented in the previous section can be conceptually inverted. A main result in section 9 was expressed in eq. (26), which we can now write as

$$\beta\left(P(a^{1/2})\,x\right) = \beta(a) + \beta(x); \tag{39}$$

that is, under the change of background illumination $x \to P(a^{1/2})\,x$, the brightness of each perceived light is increased by the brightness of the perceived light a which determines the change of background illumination.

Let us consider the following weakened version of (39) as an axiom:

Axiom 7: The brightness $\beta(x)$ of $x \in \mathscr{P}$ relative to the unit $c \in \mathfrak{A}$ is a differentiable function such that

$$\beta\left(P(a^{1/2})\,x\right) = \beta(x) + \chi(a) \tag{40}$$

where $\chi(a)$ depends only on the change of background illumination.

If $x = c$, then $\beta(x) = 0$ since c is the reference standard for brightness. Also, by (39), $P(a^{1/2})\,c = a$ whence

$$\beta(a) = \beta\left(P(a^{1/2})\,c\right) = 0 + \chi(a);$$

thus (40) is equivalent to (39). Now introduce $\varphi(x) = e^{\beta(x)}$. Then (40) (or (39)) is equivalent to

$$\varphi\left(P(a^{1/2})\,x\right) = \varphi(a)\,\varphi(x).$$

Let $a = \alpha c$, $x = \beta c$ with $\alpha, \beta \in \boldsymbol{R}^+$ and find

$$\varphi\left(P(a^{1/2})\,x\right) = \varphi(\alpha\beta c) = \varphi(\alpha c)\,\varphi(\beta c). \tag{41}$$

If we write $h(\alpha) = \varphi(\alpha c)$, then (41) asserts $h(\alpha\beta) = h(\alpha)\,h(\beta)$: h is a differentiable homomorphism of $\boldsymbol{R}^+$ into $\boldsymbol{R}$, whence $h(\alpha) = \alpha^\lambda$ for some $\lambda \in \boldsymbol{R}$. Thus $\varphi(\alpha c) = \alpha^\lambda$ and therefore

$$\varphi(\alpha x) = \varphi\left(P(\alpha^{1/2}\,c)\,x\right) = \varphi(\alpha c)\,\varphi(x) = \alpha^\lambda\,\varphi(x);$$

φ is a homogeneous function of order λ. The argument given by Koecher (1962) can now be applied to conclude that

$$\varphi(x) = |x|^{\lambda/r}, \quad x \in \exp \mathfrak{H}(r, R), \tag{42}$$

where we have made use of $\varphi(c) = e^{\beta(c)} = 1$. Using the isomorphism $a \mapsto \alpha a$, we find

$$\varphi(x) = |x^{\alpha}|^{\lambda}. \tag{43}$$

Combining (34), (42), and (43), we find

$$\beta(x) = \begin{cases} \frac{\lambda}{2} \log |x| & \text{for } x \in \exp \mathfrak{H}(2, R), \\ \lambda \log (x_1^{\alpha_1} x_2^{\alpha_2} x_3^{\alpha_3}) & \text{for } (x_1, x_2, x_3) \in R^+ \times R^+ \times R^+. \end{cases} \tag{44}$$

which retrieves the brightness formulae (22) and (23) of section 9, with the unit of brightness scale factor $\varkappa = \frac{\lambda}{2}$ in the first case, and $\varkappa = \lambda$ in the second.

11. Hue and Saturation

The previous section showed that the Jordan algebra $\mathfrak{A}_{\alpha_1 \alpha_2 \alpha_3}$ which underlies Stiles' color perception model is isomorphic to $\mathfrak{R} = R \oplus R \oplus R$ with the usual field structure on each summand. Hence there is no loss of generality in limiting our analysis of the Stiles'-like models to the Helmholtz case $\alpha_1 = \alpha_2 = \alpha_3$, that is, to $\mathfrak{R}$, since the isomorphisms ψ_{α_k} enable us to pass from the Helmholtz model to the Stiles model. Introduce the reduced norm $|x| = x_1 x_2 x_3$ on $\mathfrak{R}$. If $\mathfrak{A}$ denotes any compact real Jordan algebra, then exp: $\mathfrak{A} \to \exp \mathfrak{A}$ is a bijective mapping, so the inverse map, which will be denoted log, is well defined. Let $\mathcal{N} = \{x \in \exp \mathfrak{A} : |x| = 1\}$; $\mathcal{N}$ is the *norm surface* in $\exp \mathfrak{A}$. Then $\exp \mathfrak{A} = R^+ \times \mathcal{N}$ as a product of homogeneous spaces, and $c \in \mathcal{N}$. If follows that $\log \mathcal{N}$ is a hypersurface in $\mathfrak{A}$ which passes through 0. For $\mathfrak{A} = \mathfrak{R}$, the invariant metric is

$$ds^2 = \sum \left(\frac{dx_k}{x_k} \right)^2; \tag{45}$$

if $u_k = \log x_k$, then $u = (u_1, u_2, u_3) = \log x$, $x = (x_1, x_2, x_3) \in \exp \mathfrak{R}$, so the metric induced on $\mathfrak{R}$ by the metric (45) on $\exp \mathfrak{R}$ is $\sum (du_k)^2$, the Euclidean metric. Moreover, $\log \mathcal{N} = \{v \in \mathfrak{R} : |\exp v| = 1\} = \{u \in \mathfrak{R} : \sigma(v) = 0\} = \{v \in \mathfrak{R} : v_1 + v_2 + v_3 = 0\}$, a plane in $\mathfrak{R} = R^3$ which contains 0. Consequently the restriction to $\log \mathcal{N}$ of the pullback of the metric (45) to $\mathfrak{R}$ is the Euclidean metric on $\log \mathcal{N}$, and therefore $\mathcal{N}$ itself is isometric to two-dimensional Euclidean space. We will use this fact to introduce the concepts of *hue* and *saturation*. Let $x \in \exp \mathfrak{R}$. Then we can represent x uniquely as $x = (\xi, u)$ with $\xi \in R^+$ and $u \in \mathcal{N}$. As earlier, $\beta(x) = \varkappa \log \xi$ relative to the unit $c = (1, 1, 1) \in \mathfrak{R}$. Pull u back to the plane $\log \mathcal{N}$ in $\mathfrak{R}$ by sending $u \mapsto \log u$. Introduce polar coordinates (ρ, θ) on $\log \mathcal{N}$ with ρ measuring Euclidean distance from $0 \in \mathfrak{R}$. Then, after the choice of a direction on $\log \mathcal{N}$ which determines the ray $\theta = 0$, and an orientation for $\log \mathcal{N}$, $\log u$ will have unique coordinates (ρ, θ). Define the *hue* of $x = (\xi, u) \in \exp \mathfrak{R}$ relative to $c \in \mathfrak{R}$ to be the angle θ of $\log u$. The *saturation* of x relative to c is the radial distance ρ; since log is an isometry of $\mathcal{N}$ onto $\log \mathcal{N}$, the saturation of x relative to c is just $d(c, u)$ on $\mathcal{N}$, where d is the restriction of invariant metric to the norm surface $\mathcal{N}$.

If $\mathfrak{A} = \mathfrak{H}(2, \boldsymbol{R})$, then $\mathcal{N} = \{x \in \exp \mathfrak{H}(2, \boldsymbol{R}) : |x| = x_1 x_2 - x_3^2 = 1\}$. We have shown that $\mathcal{N}$ is isometric to the upper half plane of complex numbers with the Poincaré metric, or, equivalently, to the unit disk $\mathcal{D} = \{z \in \mathbf{C} : z\bar{z} < 1\}$ with metric

$$ds^2 = \frac{dz d\bar{z}}{(1 - z\bar{z})^2}, \tag{46}$$

because the map $w \mapsto \frac{w-i}{w+i} = z$ carries the upper half plane onto the disk and takes the Poincaré metric $\frac{dw d\bar{w}}{(\operatorname{Im} w)^2}$ onto (46). One checks that the unit matrix $c = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ is carried to $0 \in \mathbf{C}$. Introduce polar coordinates on $\mathcal{D}$ centered at 0. If $x \in \exp \mathfrak{H}(2, \boldsymbol{R})$, then $x = (\xi, u)$ with $\xi \in \boldsymbol{R}^+$ and $u \in \mathcal{N}$. Carry u to a corresponding point w by composing the map from $\mathcal{N}$ to the upper half plane with the map from the upper half plane to $\mathcal{D}$. These maps are both isometries. Let the polar coordinates of w in $\mathcal{D}$ be (ρ, θ) and define the *hue* of x to be the angle θ, and the *saturation* of x relative to c to be the distance $d(c, u)$ on $\mathcal{N}$ in the restriction of the invariant metric to $\mathcal{N}$.

12. Analysis of General Visual Scenes

Previous sections have been concerned with one or several small patches of light perceived against a uniform background which fills the remainder of the visual field. The purpose of this section is to pass to the analysis of general visual scenes. The principal problem is to define the perceived light which the observer of a fixed scene S will consider as the average perceived light; for ordinary scenes and daylight vision, this average plays the role of a white standard for the observer (cp. the discussion in chapter 2 of the Riggs *et al.* fixated vision experiment). It will not do to use Newton's idea of a simple weighted average with the weights corresponding to that fraction of the visual field occupied by a given light, for the space of perceived colors does not have a Euclidean metric (although it may be isometric to a Euclidean space). The perceptual metric must play a role in determining the weights.

In order to simplify the discussion, we will assume that a given scene S exhibits only a finite number of perceived lights $x_1, x_2, \ldots, x_n \in \mathcal{P}$. Let μ_k be the ratio of the area of the scene which exhibits light x_k to the total area of the scene; then $\sum \mu_k = 1$. Let $d(x, y)$ denote the distance between x and y in a $GL(\mathcal{P})$-invariant metric. We define the average color $\bar{x}(S)$ of the scene S to be that $\bar{x} \in \mathcal{P}$ such that

$$\sum_{k=1}^{n} \mu_k \, d(x_k, \bar{x})^2 \tag{47}$$

is a minimum. Were $\mathscr{P}$ Euclidean space and d the Euclidean metric, then $\bar{x}$ would be the "center of mass" of the distribution of "mass points" x_k with respective "mass" μ_k.

If there are just two lights and $\mu_1 = \mu_2$, then $\bar{x}$ is the midpoint of the geodesic arc which connects x_1 and x_2. It is easily seen that $\bar{x}$ exists and is unique because $\mathscr{P}$ has nowhere positive curvature.

If the lights which constitute a scene are finite in number and differ only in intensity, then any two of them are proportional, so we may put $x_k = \alpha_k x$ for some $x \in \mathscr{P}$. Then $\sum \mu_k d(x_k, \bar{x})^2 = \sum \mu_k d(\alpha_k x, \bar{x})^2$ is a minimum when $\bar{x} = \bar{\alpha} x$ for some $\bar{\alpha} \in \boldsymbol{R}^+$ (by application of the triangle inequality), and then the minimizing value of $\bar{\alpha}$ is found as follows: if

$$\Phi = \sum \mu_k d(\alpha_k x, \bar{\alpha} x)^2 = \sum_k \mu_k x^2 \log^2 \left(\frac{\bar{\alpha}}{\alpha_k}\right),$$

then

$$0 = \frac{d\Phi}{d\bar{\alpha}} = \frac{1}{\bar{\alpha}} \sum_k 2 \mu_k \log \left(\frac{\bar{\alpha}}{\alpha_k}\right),$$

whence

$$\log \bar{\alpha} \sum_k \mu_k = \sum_k \mu_k \log \alpha_k;$$

from $\sum \mu_k = 1$, conclude

$$\log \bar{\alpha} = \sum \mu_k \log \alpha_k;$$

the brightness of the average color $\bar{x}(S)$ is the weighted average of the brightness of the lights which constitute the scene.
The significance of the average color $\bar{x}(S)$ of the scene S is that color adaption appears to occur relative to $\bar{x}(S)$, and in this sense, the observer treats $\bar{x}(S)$ as a private standard "white" light (cp. Land and McCann 1971). We formalize this as

Axiom 8: If S is a visual scene, then $\bar{x}(S)$ is the observer's standard white light.

If S consists of two small light patches x_1, x_2 viewed against a uniform background b, then the weights μ_1 and μ_2 will be close to 0 and $\mu_b \cong 1$, from which one concludes that $\bar{x} \cong b$; the background illumination will constitute the observer's standard white light.

According to Axiom 8, an observer identifies $\bar{x}$ as standard white and therefore also identifies this point with the unit c of the Jordan algebra $\mathfrak{A}$ associated with $\mathscr{P}$. This enables the observer to classify perceived lights according to their brightness, hue, and saturation as defined in section 11 and therewith provides the link between theory and experimental measurements.

13. Geodesics and Complementary Colors

The geodesics through $c \in \exp \mathfrak{A}$ are of the form

$$x(\tau) = \exp \tau a \tag{48}$$

where a is the tangent vector of length 1 to the geodesic at $x=c$ (the length condition is equivalent to $\sigma(a^2)=1$ where σ denotes the reduced trace), and τ is the measure of arc length in the invariant metric. The distance τ from c to $x(\tau)$ can be explicitly found as follows: from (48),

$$\log x(\tau)=\tau a,$$

so

$$\sigma(\log^2 x(\tau))=\sigma((\tau a)^2)=\tau^2\sigma(a^2)=\tau^2,$$

whence

$$d(c,x)=\sigma^{1/2}(\log^2 x). \tag{49}$$

Now join $x\in\mathscr{P}$ to c by the geodesic $x(\tau)=\exp\tau a$; then the continuation of the geodesic past c in the direction from x to c corresponds to the curve $x(-\tau)=\exp -\tau a$. Let x^* denote that point on this geodesic arc such that $d(c,x^*)=d(c,x)$. If the coordinates of x with respect to the decomposition $\mathscr{P}=\boldsymbol{R}^+\times\mathscr{N}$ are $x=(\xi,u)$, then $x^*=(\xi^{-1},u^{-1})$ where u^{-1} denotes the inverse of u considered as an element of the Jordan algebra $\mathfrak{A}$. The Jordan inverse, $x\mapsto x^{-1}$, is an isometry of any $GL(\exp\mathfrak{A})$-invariant metric. Indeed, $dx^{-1}=$ $=-P^{-1}(x)\,dx$ so, if $y=x^{-1}$, then

$$\begin{aligned} ds^2&=\sigma((P^{-1}(y)\,dy)\,dy)=\sigma((P^{-1}(x^{-1})\,dx^{-1})\,dx^{-1})\\ &=\sigma((P(x)\,P^{-1}(x)\,dx)\,(P^{-1}(x)\,dx))\\ &\sigma(dx\,(P^{-1}(x)\,dx))=\sigma((P^{-1}(x)\,dx)\,dx),\end{aligned}$$

which proves the invariance. It follows that $d(c,u)=d(c,u^{-1})$: in other words, x and x^* are equally saturated; and $\log u^{-1}=-\log u$, so the hue of x^* is the *complement* of the hue of x. Finally, since $(\exp\tau a)^{-1}=\exp -\tau a$, it follows that $x^*=x^{-1}$, which demonstrates that x and x^{-1} are perceived colors which have the same saturation, complementary hue, and complementary brightness relative to c; the latter means

$$\beta(x^{-1})=-\beta(x).$$

We speculate that the interpretation of the involutive map $x\mapsto x^{-1}$ is simply this: it is the map which makes an originally photographed scene correspond to a photographic color negative of the scene. Further application of $x\mapsto x^{-1}$ corresponds to conversion of the negative into a photographic positive print of the scene.

14. An Example

Suppose three rooms arranged as shown in Fig. 1. The wall common to each pair of rooms is pierced by a small regularly shaped hole which appears as a small patch of color to observers 1 and 2. We suppose each room illuminated by various distinct white lights, which are perceived as w_E, w_1, w_2 by E, 1, 2 respectively. The observer E plays the role of experimenter, whose task is to

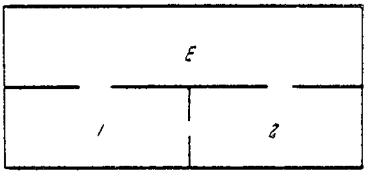

Fig. 1

determine the relationships amongst the background and light admitted through the hole as seen by the various observers. For the sake of simplicity, we will suppose that w_E, w_1, and w_2 are all of the same brightness; then without loss of generality, we may assume that all three perceived lights lie on $\mathcal{N}$.

If $\mathcal{P} \cong R^+ \times SL(2, R)/SO(2)$, then $\mathcal{N}$ is isometric to the unit disk $\mathcal{D}$ with the Poincaré metric. The geodesics of this metric are arcs of circles orthogonal to the unit circle (understood to include segments of diameters as well). The situation is illustrated by Fig. 2, in which the experimenter E has placed his perceived color w_E at the origin of $\mathcal{D}$, that is, E has identified w_E with standard white, according to Axiom 8; consequently the hue $H_{1/E}$ of 1 (resp., $H_{2/E}$ of 2) as seen by E is measured by the angle between the positive X-axis and the radius from E through 1 (resp., 2), and the saturation of 1 (resp., 2) relative to E is the Poincaré distance of w_1 (resp., w_2) from E.

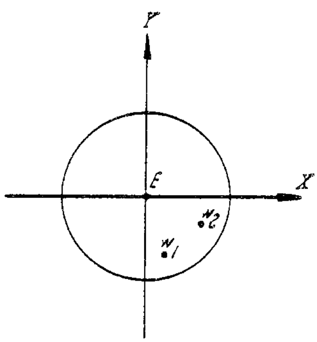

Fig. 2

Each of 1, 2 will consider their respective background illuminant w_1, w_2 to constitute the standard white illuminant, and hence each considers his illuminant to lie at E in the homogeneous space $\mathscr{D}$. Since geodesics are carried onto geodesics by isometries of $\mathscr{D}$, it follows that w_1 views w_2 along the geodesic arc which joins these points, and that w_2 views w_1 along the same arc, but in the reverse direction. This situation is diagrammed in Fig. 3.

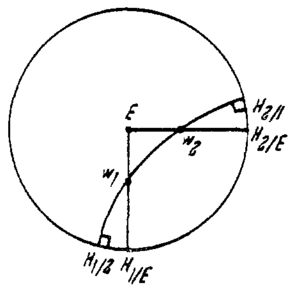

Fig. 3

The hue $H_{2/1}$ of 2 seen by 1 and the hue $H_{1/2}$ of 1 seen by 2 differ more than the hues $H_{1/E}$ and $H_{2/E}$ seen by E. Moreover, if either w_1 or w_2 is transported to the origin (the current position of E) by an isometry, then the geodesic joining w_1 and w_2 will be mapped onto a diameter, whence $H_{1/2}$ and $H_{2/1}$ are seen to be perceived as complementary colors by both w_1 and w_2, but not by E.

Finally, we notice that the three observers enjoy perfect symmetry in the arrangement; none can claim special perceptual authority. From this it follows that the statements made above must hold for any permutation of the designating symbols E, 1, and 2 (cp. Yilmaz 1962).

The analysis is somewhat more complicated if the three perceived lights are not of the same brightness, but the general principles remain the same.

An analogous argument can be made if $\mathscr{P} = \boldsymbol{R}^+ \times \boldsymbol{R}^+ \times \boldsymbol{R}^+$, but in this case the resulting surface $\mathscr{N}$ is isometric to Euclidean space, which somewhat simplifies the geometry.

15. A Critical Experiment

It is of interest to determine which of the models of $\mathscr{P}$ provides the more accurate representation of the phenomena of color perception. The purpose of this final section is to propose a decisive experiment for making this distinction.

Suppose x, y, z are three colors of the same brightness. We may assume without loss of generality that $x, y, z, \in \mathscr{N}$. Moreover, as a model of $\mathscr{N}$ we may choose the Poincaré disk $\mathscr{D}$, or the Euclidean plane R^2, the latter only after the introduction of appropriate isomorphisms ψ_{z_k} of the Jordan algebra R in case of Stiles' model. We propose to use the distinct properties of Euclidean spaces and spaces of constant negative curvature to predict perceptual phenomena whose presence will confirm, and whose absence will deny, the validity of the particular geometry considered.

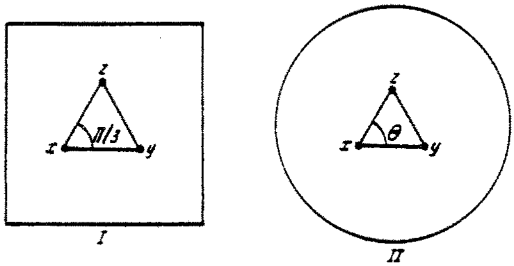

Fig. 4

Let x, y, z be vertices of an equilateral triangle in each case, as shown in Fig. 4, and suppose that the distances $d(x, y)$, $d(x, z)$, $d(y, z)$ are increased without an upper bound while the conditions $d(x, y) = d(y, z) = d(z, x)$ are maintained in each case. For the Euclidean case I, observer x will continue to see y and z differing by a hue of angle $\pi/3$, whereas for the non-Euclidean case II, as the sides of the triangle increase, the angles will approach zero, and x will observe the convergence of the hues of y and z. These conclusions reflect intrinsic properties of the two geometries and should provide an experimental method for excluding one of them from consideration as a model of the space of perceived colors.

16. Summary

In conclusion we summarize the principal results of this paper.

When combined with well known facts about the affine structure of color space $\mathscr{P}$, the hypothesis of local homogeneity with respect to changes of background illumination is equivalent to the hypothesis of global homogeneity. There are two distinct types of homogeneous space compatible with the affine assumptions: $\mathscr{P} = R^+ \times R^+ \times R^+$, and $\mathscr{P} = R^+ \times SL(2, R)/SO(2)$.

If $\mathscr{P}$ is a homogeneous space of the group G, then a color metric on $\mathscr{P}$ must be G-invariant, and this condition determines the color metric up to selection of the unit of measurement on each irreducible factor of $\mathscr{P}$. If $\mathscr{P} = \boldsymbol{R}^+ \times \boldsymbol{R}^+ \times \boldsymbol{R}^+$, the resulting metric is Stiles' generalization of Helmholtz' metric, and the resulting color space is isometric to Euclidean space. If $\mathscr{P} = \boldsymbol{R}^+ \times SL(2, \boldsymbol{R})/SO(2)$, the metric appears to be new, and the resulting color space is not isometric to Euclidean space.

The notions of brightness, hue, and saturation have natural interpretations as point-pair invariant functions, which is consistent with and provides an explanation of well known invariance properties of relative perception of colors.

Finally, an experiment is proposed which will discriminate between the two types of color space geometries.

Acknowledgement

The author expresses his appreciation to Professor Harry Sperling of the Graduate School of Biomedical Science, University of Texas at Houston, for numerous valuable discussions, and is also pleased to acknowledge the catalytic effect of Yilmaz' interesting paper (Yilmaz 1962).

References

[1] Abney, W. de W.: Researches in Colour Vision. London: Longmans, Green 1913.
[2] Abney, W. de W., Festung, E. R.: Colour photometry. Phil. Trans. Royal Soc. London *177*, 423—456 (1886).
[3] Braun, H., Koecher, M.: Jordan Algebren. Berlin-Heidelberg-New York: Springer 1966.
[4] Cornsweet, T. N.: Visual Perception. New York: Academic Press 1970.
[5] Cornsweet, T. N., Pinsker, H. M.: Luminance discrimination of brief flashes under various conditions of adaptation. J. Physiology *176*, 294—310 (1965).
[6] Descartes, René: Météores 1638.
[7] Ekman, Gösta: Weber's Law and related functions. J. Psych. *47*, 343—352 (1959).
[8] Euler, Leonhard: Essai d'une explication physique des couleurs engendrées sur des surfaces extrêmement nunces. Mém. de l'académie des sci. de Berlin *8*, 262—282 = Opera Omnia, Ser. (*3*), *5*, No. 209, 156—171 (1752).
[9] Fechner, G. T.: Sachen der Psychophysik. Leipzig 1877.
[10] Goethe, Johann Wolfgang von: Theory of Colours, 1810. (English trans., reprinted.) Cambridge: The MIT Press.
[11] Graham, Clarence H., ed.: Vision and Visual Perception. New York: J. Wiley 1965.
[12] Grassmann, Hermann Günter: Zur Theorie der Farbenmischung. Ann. Physik und Chemie *89*, 69—84 (1853) = Phil. Mag. *7*, 254—284 (1853).
[13] Hecht, S.: A theory of visual discrimination. J. General Physiology *18*, 767—789 (1935).
[14] Helgason, S.: Differential Geometry and Symmetric Spaces. New York: Academic Press 1962.
[15] Helmholtz, Hermann von: Über die Theorie der zusammengesetzten Farben. Ann. Physik und Chemie *87*, 45—66 (1852).
[16] Helmholtz, Hermann von: Treatise on Physiological Optics. 1866. English trans. of first German edition, Optical Society of America, 1924.
[17] Helmholtz, Hermann von: On Goethe's scientific researches. 1876a. Popular Lectures, English trans. London: Longmans, Green and Co. 1908.
[18] Helmholtz, Hermann von: On the relation of optics to painting. 1876b. Popular Lectures on Scientific Subjects, second series, English trans. London: Longmans, Green and Co. 1908.
[19] Helmholtz, Hermann von: Versuch einer erweiterten Anwendung des Fechnerschen Gesetzes im Farbensystem. Z. Psychol. Physiol. Sinnesorg. *2*, 1—30 (1891).
[20] Helmholtz, Hermann von: Versuch, das psychophysische Gesetz auf die Farbenunterschiede trichromatischer Augen anzuwenden. Z. Psychol. Physiol. Sinnesorg. *3*, 1—20 (1892a).

[21] Helmholtz, Hermann von: Kürzeste Linien im Farbensystem. Z. Psychol. Physiol. Sinnesorg. *4*, 108—122 (1892b).
[22] Hering, E.: Zur Lehre vom Lichtsinne. Wien: Carl Gerold's Sohn 1878.
[23] Hooke, Robert: Micrographia, Reprint of edition of 1665. New York: Dover Pub. 1961.
[24] Jacobson, N.: Lie Algebras. New York: Interscience Pub. 1962.
[25] Koecher, M.: Jordan Algebras and their Applications. Lecture notes, Univ. of Minnesota, 1962.
[26] Land, Edwin, McCann, John J.: Lightness and retinex theory. J. Opt. Soc. Amer. *61*, 1—11 (1971).
[27] MacAdam, David L.: Color discrimination and the influence of color contrast on visual acuity. Rev. d'Optique *28*, 161—173 (1949).
[28] MacAdam, David L.: Loci of constant hue and brightness determined with various surround colors. J. Opt. Soc. Amer. *40*, 589—595 (1950).
[29] MacAdam, David L.: Sources of Color Science. Cambridge: The MIT Press 1970.
[30] Mach, Ernst: The Principles of Physical Optics. 1913. Reprint of English trans. New York: Dover Pub.
[31] Maxwell, James Clerk: Theory of the perception of colors. Trans. Royal Scottish Soc. Arts *4*, 394—400 (1856).
[32] Maxwell, James Clerk: The diagram of colors. Trans. Royal Soc. Edinburgh *21*, 275—298 (1857).
[33] Maxwell, James Clerk: Theory of compound colors, and the relations of the colors of the spectrum. Proc. Royal Soc. London *10*, 404—409 (1860).
[34] Maxwell, James Clerk: On color vision. Proc. Royal Inst. Great Britain *6*, 260—271 (1872).
[35] Newton, Isaac: New theory about light and colors. Phil. Trans. Royal Soc. London *80*, 3075—3087 (1671).
[36] Newton, Isaac: Opticks. London: Smith and Walford 1704.
[37] Resnikoff, H. L.: A generalization of the psychophysical laws of Fechner and Stevens, to appear.
[38] Riemann, Bernhard: Über die Hypothesen, welche der Geometrie zu Grunde liegen. 1854, in: Werke, reprint. New York: Dover Pub. 1953.
[39] Riggs, L. A., Ratliff, F., Cornsweet, J. C., Cornsweet, T. N.: The disappearance of steadily fixated test objects. J. Opt. Soc. Amer. *43*, 495—501 (1953).
[40] Schier, Donald S.: Louis Bertrand Castel, Anti-Newtonian Scientist. Cedar Rapids: The Torch Press 1941.
[41] Schrödinger, E.: Grundlinien einer Theorie der Farbenmetrik in Tagessehen. Ann. Physik *63*, 397—426, 427—456, 481—520 (1920).
[42] Stevens, S. S.: On the psychophysical law. Psych. Rev. *64*, 153—181 (1957).
[43] Stevens, S. S.: Neural events and the psychophysical law. Science *170*, 1043—1050 (1970).
[44] Stiles, W. S.: A modified Helmholtz line-element in brightness-colour space. Proc. Phys. Soc. London *18*, 41—65 (1946).
[45] Stiles, W. S., Wyszecki, Günter: Color Science. New York: J. Wiley and Sons 1967.
[46] Vos, J. J., Walraven, P. L.: An analytical description of the line element in the zone-fluctuation model of colour vision. — I. Basic concepts. — II. Derivation of the line element. Vision Research *12*, 1327—1344, 1345—1365 (1972).
[47] Whittaker, E. T.: A History of the Theories of Aether and Electricity, Vol. I: Classical Theories. 1910. London: Thomas Nelson and Sons, revised 1951.
[48] Yilmaz, H.: On color perception, in: Evidence for Gravitational Theories, pp. 239—251. New York: Academic Press 1962.
[49] Young, Thomas: On the theory of light and colors. Phil. Trans. Royal Soc. London *92*, 20—71 (1802).
[50] Young, Thomas: A Course of Lectures on Natural Philosophy and the Mechanical Arts. London: Taylor and Walton 1845.

Dr. Howard L. Resnikoff
Department of Mathematics
Rice University
Houston, TX 77001, U.S.A.