

## **Geometry in Color Perception**

**Abhay Ashtekar  
Alejandro Corichi  
Monica Pierri**

Vienna, Preprint ESI 517 (1997)

December 22, 1997

Supported by Federal Ministry of Science and Research, Austria  
Available via <http://www.esi.ac.at>

# Geometry in Color Perception

Abhay Ashtekar<sup>1,3</sup> Alejandro Corichi<sup>1,2</sup> and Monica Pierr<sup>1</sup>

*1 Center for Gravitational Physics and Geometry*

*Department of Physics, Penn State,*

*University Park, PA 16802-6300, USA*

*2 Instituto de Ciencias Nucleares*

*Universidad Nacional Autónoma de México*

*A. Postal 70-543, México D.F. 04510, México*

*3 Erwin Schrödinger International Institute for Mathematical Sciences*

*Boltzmannngasse 9, A-1090 Vienna, Austria*

## Abstract

Since the days of Newton, the space  $\mathcal{C}$  of colors perceived by a normal human eye has been known to be three-dimensional. It was recognized as one of the earliest examples of a manifold by Riemann, and its structure was further analyzed by Maxwell, Helmholtz, Schrödinger and others. In this note we give an elementary introduction to the geometry of this space. In the spirit of mathematical physics, we construct a simple model of color perception and derive from it qualitative features of  $\mathcal{C}$ . The model has an important limitation; it ignores brain interventions which can be important. The discussion is aimed at advanced undergraduate or beginning post-graduate students.

Typeset using REVTeX

## I. INTRODUCTION

Of the five human senses, only two can be excited essentially instantaneously by signals from distant sources: hearing and sight. In both cases, signals propagate as waves. It is not an accident therefore that we have television but not its equivalent in the realm of smell, taste or touch. Smell for example, propagates by diffusion, a slow process which is difficult to control, while taste and touch require immediate contact with the source of the signal. There is, however, a rather fundamental difference also between hearing and sight. Hearing occurs because sound excites cells in the cochlear tube of the inner ear and each little longitudinal section of the tube is sensitive to a different frequency. Consequently, we are capable of distinguishing a very large class of ‘sound profiles’. Thus, the space of sounds that human beings are capable of hearing has a very large dimension, which, for all practical purposes, can be taken to be *infinite*. This is in striking contrast with sight. It turns out that the space of colors that human beings are capable of seeing is only *three* dimensional!

The subject of color has a long and distinguished history and some of the most creative mathematicians, physicists and philosophers have devoted years of study to unravel its mysteries.<sup>1</sup>As is the case with so many fascinating discoveries, here too the road begins with Newton [1]. It was he who first realized that distinct intensity (or spectral) distributions  $I(\lambda)$  of light are often indistinguishable to the human eye. In the modern terminology, he concluded that a *perceived color* is an *equivalence class*  $\{I(\lambda)\}$  of spectral distributions and then went on to argue that, while the space of all spectral distributions  $I(\lambda)$  is infinite dimensional, the space  $\mathcal{C}$  of perceived colors  $\{I(\lambda)\}$  is only three-dimensional. A century later, in 1802 Thomas Young [2] made the bold conjecture that the origin of this three-dimensionality may lie in the property that color sensation is caused by absorption of light by three different pigments in the human retina. It took another century and a half to definitively confirm this conjecture experimentally [3]!

In the meantime, in 1853 Grassmann [4,5] gave a clean mathematical formulation of Newton’s ideas. The following year, Riemann gave his epoch making inaugural address in Göttingen, in which he laid down the foundations of differential geometry. During this lecture, he singled out the physical space in which we live and the space of perceived colors as possible examples of ‘curved spaces’ of *three* dimensions [6]. The first of these has been quoted often in the context of general relativity. However, perhaps it is the second that constitutes a more immediate example, as far as our common experience is concerned. Maxwell sharpened these notions and then used them in 1861 to explicitly demonstrate that color photography was possible, despite the fact that the photographic emulsions themselves were black and white. The method of his demonstration supported Young’s hypothesis. About the same time, Helmholtz was interested in our perception of both sound and color and made significant contributions to both areas. In particular, in 1891, he suggested [7] a form of the metric on the space  $\mathcal{C}$  of perceived colors which determines the ‘distance’ between any two

---

<sup>1</sup>For example, the chemist John Dalton, the poet Johann Wolfgang von Goethe, and philosophers John Locke and Arthur Schopenhauer all had theories of color of their own. Ludwig Wittgenstein’s last work was *Remarks on Colour*.

colors, i.e., how far or close an average human subject perceives the two colors to be. A little later, Schrödinger devoted several years to the study of colors. Indeed, this was one of his primary scientific preoccupations during his Vienna years. In 1920, he proposed another, ‘improved’ form of the metric [8]. Since then, there have been several other attempts at improvements, perhaps the most notable being that of Stiles [9].

The aim of this contribution is to present a simple physical model of color perception and derive from it the geometric structure of  $\mathcal{C}$  in a systematic, pedagogical fashion. In the spirit of this volume, the account is addressed to advanced undergraduate or beginning post-graduate students and assumes only a rudimentary knowledge of differential geometry.

Historically, the basic, geometric structure of  $\mathcal{C}$  was discovered slowly, using simple experiments on human subjects. In the spirit of mathematical physics, one can therefore ask if this well-known structure can be derived systematically from first principles, knowing only adsorption coefficients of cones in the human eye that are now known. We will see that the answer is in the affirmative. This derivation is a simple example of a *phenomenological* application of differential geometry; one which is qualitatively similar to but *significantly* simpler than the more well-known applications to, e.g., cosmology where one uses the observed large scale homogeneity of the universe to severely constrain its geometry. We should emphasize that this is *not* meant to be a scholarly account. In particular, we will use terminology which should be intuitively obvious to mathematical physicists, even though it occasionally differs from that used in the scholarly literature, decreed by the International Commission on Illumination. (For example, what we call ‘perceived color’ is called ‘psycho-physical color’ in the standard lore.)

The structure of the paper is as follows. Section II contains the main mathematical discussion. The model we discuss is very simple, based only on physical –rather than psycho-physical– aspects of color perception. It ignores the intervention of the brain which in fact plays an important role in certain aspects of color perception. These limitations are discussed in section III.

## II. GEOMETRY

This section is divided in to three parts. In the first, we outline the procedure by which early researchers in the field arrived at a qualitative picture (figure 1) of the space  $\mathcal{C}$  of perceived colors. In the second part, we will show how one can derive this structure in a straightforward fashion starting from the qualitative features of the experimentally observed adsorption curves. In the third, we discuss the Riemannian geometry of this space.

### A. Phenomenology

The key experimental technique used to unravel the structure of the space of perceived colors was quite simple: Mix and match. More precisely, mixtures of spectral distributions  $I(\lambda)$  were exposed to the a large number of human subjects who were asked to comment on whether they perceived the same color or different, and, if different, how ‘close’ or ‘far’ these colors seemed to be. From the resulting data, qualitative features of the space  $\mathcal{C}$  of perceived colors could be deduced.

First, let us consider the so-called *spectral colors*, i.e., those spectral distributions which can *not* be further decomposed by a prism or a grating. In the idealized description, in this case, we have

$$I(\lambda) = k\delta(\lambda, \lambda_0) \quad (2.1)$$

for some constant  $k$  and a fixed wave-length  $\lambda_0$ , and, in a more realistic description,

$$I(\lambda) = k \begin{cases} \frac{1}{\epsilon}, & \text{if } \lambda_0 - \frac{\epsilon}{2} < \lambda < \lambda_0 + \frac{\epsilon}{2} \\ 0, & \text{otherwise} \end{cases} \quad (2.2)$$

for some  $\epsilon \ll 1$ . It is an observed fact that each spectral color is perceived as distinct and a mixture of two spectral colors produces a new color; spectral colors can not be mixed to obtain another spectral color. Furthermore, binary mixtures of spectral colors suffice to produce all observed colors. That is, any color produced by mixing three or more spectral colors can also be produced by mixing a (different, in general) pair of spectral colors. Finally, the space of spectral colors can be parametrized by two numbers, the wave-length  $\lambda_0$  which takes values in the visual range and the intensity  $k$ . It is thus a two-dimensional space.

The second important feature of  $\mathcal{C}$  is that it has an underlying convex-linear structure. To see this, let us first introduce some notation. Denote by  $\mathcal{I}$  the space of all spectral distributions,

$$\mathcal{I} := \{I(\lambda) : (-1, 1) \longrightarrow R; \quad I(\lambda) \geq 0.\} \quad (2.3)$$

where for notational simplicity we have normalized the spectrum of visible light to lie in the interval  $(-1, 1)$ . Because  $I(\lambda)$  are required to be non-negative,  $\mathcal{I}$  is not a vector space. However, if  $I_n(\lambda)$  are elements of  $\mathcal{I}$ , then for all constants  $a_n \geq 0$ ,  $\sum a_n I_n(\lambda)$  also belong to  $\mathcal{I}$  if  $n$  runs over a finite range. (In particular, if  $I_1(\lambda)$  and  $I_2(\lambda)$  belong to  $\mathcal{I}$ , so does the ‘line’ joining them,  $aI_1(\lambda) + (1 - a)I_2(\lambda)$  where  $0 \leq a \leq 1$ .) That is,  $\mathcal{I}$  an infinite-dimensional, convex sub-space of the vector space consisting of all functions on the interval  $(-1, 1)$ . Since distinct spectral intensities can be perceived as the same color, there is a (in fact, huge!) projection  $P$ :

$$P : \mathcal{I} \longrightarrow \mathcal{C} \quad (2.4)$$

Let us denote by  $c$  the image of the spectral distribution  $I(\lambda)$ . There is an obvious equivalence relation  $\sim$  defined on  $\mathcal{I}$  given by:  $I(\lambda) \sim I'(\lambda) \iff P(I) = P(I')$ . Thus, we have a 1 – 1 relation between these equivalence classes and perceived colors:  $\{I(\lambda)\} \equiv c$ . Spectral distributions  $I(\lambda)$  which belong to the same equivalence class,  $c$ , are called *metamers* (of the color  $c$ ). Now, observations show that the projection map  $P$  has a key property: it preserves linearity. More precisely, for each integer  $n$ , suppose that the spectral distributions  $I_n(\lambda)$  and  $I'_n(\lambda)$  are metamers, i.e.,  $\{I_n(\lambda)\} = \{I'_n(\lambda)\}$ . Then, their linear combinations are also metamers (of a different color):

$$\{I_n(\lambda)\} = \{I'_n(\lambda)\} \implies \left\{ \sum_n a_n I_n(\lambda) \right\} = \left\{ \sum_n a_n I'_n(\lambda) \right\} \quad (2.5)$$

for all non-negative coefficients  $a_n$ . Therefore, the space  $\mathcal{C}$  of perceived colors inherits from  $\mathcal{I}$  the structure of a convex-linear space.

## FIGURES

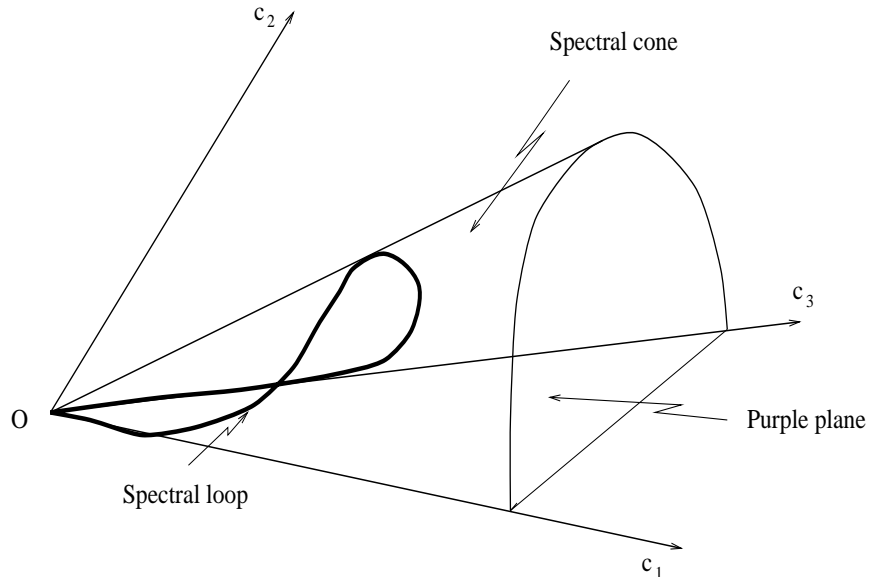


FIG. 1. The general structure of the space  $\mathcal{C}$  is shown. The Spectral cone together with the Purple plane are the boundary of the convex set  $\mathcal{C}$ . The axes are not drawn at ‘right angles’ because there is no metric to define the notion of orthogonality.

By putting together the observations of the last two paragraphs, one can conclude that the spectral colors constitute a boundary of  $\mathcal{C}$ . Since the boundary is two-dimensional, *the space  $\mathcal{C}$  is a 3-dimensional convex-linear space*. This underlying mathematical structure is reminiscent of the structure of states (i.e. density matrices) in quantum mechanics. The space of all states is also convex-linear, the boundary consists of pure states and any mixed state can be obtained by a superposition of pure states. In the present case, the spectral colors are the analogs of pure states.

Finally, let us consider the structure of the part of the boundary of  $\mathcal{C}$  formed by spectral colors. The boundary is parametrized by  $\lambda_0$ , the wave length of the spectral color and  $k$  the intensity. Hence, it is of interest to examine the one-dimensional curve made of all colors of unit intensity ( $k = 1$ ). It turns out that this curve is a closed loop; it is called the *spectral loop*.

Figure 1 summarizes the qualitative structure of  $\mathcal{C}$  that had been arrived at phenomenologically.  $\mathcal{C}$  is the manifold with boundary that resembles a solid cone. It turns out that the complete boundary consists of a *spectral cone* and the *purple plane*. As one moves along the spectral loop, one runs through distinct spectral colors while motion along ‘radial’ direction corresponds only to change in the intensity of these colors. The purple plane consists of those colors which can be obtained by combining extreme violet and extreme red.

### B. Derivation of the structure of $\mathcal{C}$

Let us now see, in the spirit of mathematical physics, how the structure of  $\mathcal{C}$  can be derived from first principles.

The experimental input will now consist only of the observed adsorption curves of the three cones in the human eye, which are qualitatively sketched in figure 2. All results of this sub-section will follow from this information.

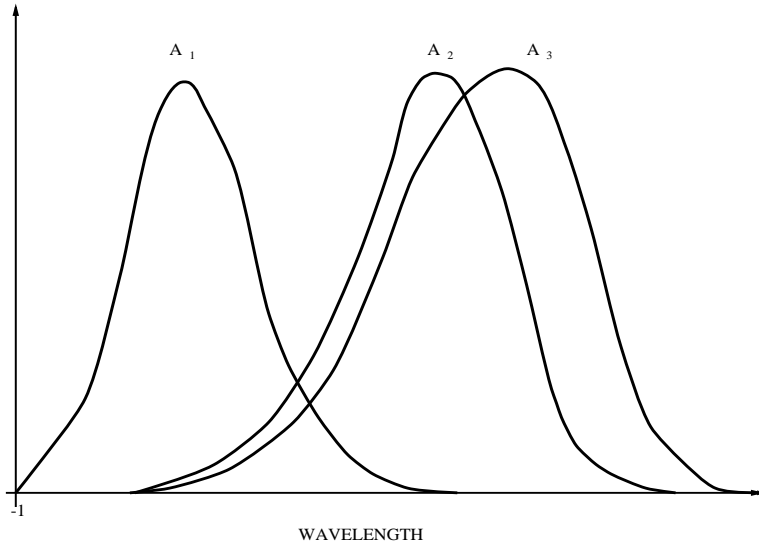


FIG. 2. The adsorption functions for the three eye cones. The actual domain, (380nm, 780nm) is normalized to the interval  $(-1, 1)$  for simplicity.

Denote the three adsorption functions by  $A_i(\lambda)$ ,  $i = 1, 2, 3$ . From physical considerations, it is obvious that, given an incident spectral distribution  $I(\lambda)$ , there are three ‘output’ signals  $c_i$ , one for each cone, given by

$$c_i := \int_{-1}^1 d\lambda I(\lambda) A_i(\lambda) \quad (2.6)$$

sent to the brain. Therefore, if we ignore potential complications in the brain itself –these are discussed in Sec. III– we can characterize each perceived color by the three numbers,  $c_i$ . Thus, in the simplest mathematical model, we have a precise formula for the projection map  $P$  (of (2.4)) which maps the infinite dimensional space of spectral distributions to the three dimensional space of perceived colors:

$$P[I(\lambda)] = c_i. \quad (2.7)$$

We will use this map to recover the structure of the space of colors depicted in figure 1, obtained in Sec II.A from several phenomenological inputs.

First of all, it is obvious that the space  $\mathcal{C}$  is three dimensional. Next, since we now have the *explicit* form of the projection map, we study its properties. It is obvious that the map is convex-linear, i.e. satisfies (2.5). Note that this is now a *derived property*, not an independent phenomenological input that it was in Sec. II A. Thus, we have shown that  $\mathcal{C}$  has a natural convex-linear structure.

Next, the space is endowed with a natural ‘radial’ vector field. To see this, note that, since the space of spectral distributions,  $\mathcal{I}$ , is naturally embedded in the vector space of all functions on the interval  $(-1, 1)$ , it has a natural radial vector field. Motions along the integral curves of this vector field correspond to spectral distributions which have the same

profile except for an overall multiplicative factor. This vector field naturally projects to a vector field  $B^a$  on  $\mathcal{C}$ . This is the *brightness vector field* and its components in the natural chart  $c_i$  are given by  $B^i = c_i$ . The integral curves of this vector field are rays through the origin and, as the name suggests, points that lie on any one ray represent colors which are all the same except for brightness.

Finally, let us look at spectral colors, i.e., (idealized) spectral distributions (2.1). Because of the brightness vector field, the structure of this sub-set can be understood entirely in terms of the structure of the spectral curve, i.e., colors corresponding to incident intensities  $I(\lambda) = \delta(\lambda, \lambda_0)$ . Since these colors correspond to coordinates  $c_i = A_i(\lambda_0)$ , the spectral curve  $\mathbf{c}(\lambda_0)$ , parametrized by  $\lambda_0$  is given simply by

$$c_i(\lambda_0) = A_i(\lambda_0) \tag{2.8}$$

in the natural chart. Some of the elementary properties of the spectral curve now follow immediately from the form of the adsorption functions (i.e., figure 2). To see this, let us embed  $\mathcal{C}$  in  $R^3$  (without a specific metric) as in figure 1. Let us begin at the origin,  $c_i = 0$ , the totally black color corresponding to the incident intensity  $I(\lambda) = 0$ , and follow the spectral curve as  $\lambda$  increases from  $-1$  to  $1$ . If we move a little bit, only the first set of cones is excited, whence only  $A_1$  is non-zero, i.e., for a finite interval, the spectral curve  $\mathbf{c}(\lambda_0)$  is tangential to the  $c_1$ -axis. Physically, this part of the spectral curve corresponds to the red end of the spectrum. As we continue to move, we reach the value of  $\lambda$  at which the function  $A_2$  starts taking values different from zero. The curve  $\mathbf{c}(\lambda_0)$  then lies in the  $(c_1, c_2)$  plane. The curve  $\mathbf{c}(\lambda_0)$  departs that plane when  $A_3$  stops being zero, and it goes in to the ‘octant’ defined by the vectors  $\partial/\partial c_i$  with positive coefficients. Thus,  $\mathbf{c}(\lambda_0)$  traces a convex curve and then hits the  $(c_2, c_3)$  plane; it then goes to the axis  $c_3$  and finally reaches back the point  $O$ . Since the final segment lies in the  $(c_2, c_3)$  plane, the spectral curve has a discontinuity at the origin. Thus, we conclude:

1. The starting point and the end point of the curve coincide:  $\mathbf{c}(\lambda_0 = -1) = \mathbf{c}(\lambda_0 = 1) = O$ . Thus, the spectral curve is indeed a loop.
2. The derivative of the curve at the origin  $O$  is discontinuous:

$$\left. \frac{d\mathbf{c}(\lambda_0)}{d\lambda_0} \right|_{-1} \neq \left. \frac{d\mathbf{c}(\lambda_0)}{d\lambda_0} \right|_1 \tag{2.9}$$

3. The two-surface generated by rays starting from  $O$  and passing through points on the spectral loop –the *spectral cone*– is a boundary of a 3-dimensional convex set.

Thus, detailed properties of the spectral curve that had previously been arrived at phenomenologically from a large number of facts can be readily derived in our model. (Incidentally, Newton did not realize that the spectral curve passes through the origin. This fact as well as the discontinuity at the origin was however well-known by Schrödinger’s time.)

Finally, let us ‘explain’ the origin of the purple plane. Consider incident intensity profiles  $I(\lambda)$  whose support is the union of the extreme red part of the spectrum (i.e., on wave-lengths where only  $A_1 \neq 0$ ) and the extreme violet (i.e., wave-lengths where only  $A_3 \neq 0$ ).

The corresponding colors lie in the  $c_1$ - $c_2$  plane and fill it. *This is the purple plane which constitutes a part of the boundary of  $\mathcal{C}$ .* Note that the  $c_2$ - $c_3$ -plane or the  $c_1$ - $c_2$  plane does not constitute an analogous boundary because the form of the overlap of the  $A_2$  adsorption curve with the  $A_1$  and the  $A_3$  curves implies that most points on these planes do not lie in the image of the map  $P$  of (2.7).

Let us summarize. The basic assumption of the model is simply that the output signal from the retina to the brain be given by the overlap integrals of the incident intensity and the three adsorption functions. This assumption is well-motivated since the spectral distribution  $I(\lambda)$  represents the flux of energy per unit wave-length, per unit solid angle that is incident on the eye and the adsorption coefficients  $A_i(\lambda)$  physically represent the fraction of the incident energy that is absorbed by the  $i$ th set of cones at the wavelength  $\lambda$ . The model is ‘compact’ in the sense that this is the *only* assumption. From it, one can systematically derive that the structure of the space of perceived colors is that of figure 1, i.e., is the one that had been put together slowly from diverse phenomenological inputs. The model is of course a drastically simplified version of reality since it ignores the role played by the rods in the retina and, more importantly, mechanisms in the brain. However, such idealizations are in the best tradition of theoretical physics and it is pleasing that the model is capable of providing complete insight in to the qualitative features of the manifold structure of  $\mathcal{C}$ .

### C. Riemannian Structure

We can now turn to the next key question: that of finding a physical metric on the space  $\mathcal{C}$ . We can all tell if the two colors are close to one another or far apart and the mathematical physics problem is that of introducing a positive definite metric  $g_{ab}$  on  $\mathcal{C}$  which correctly captures this perception. Thus, what we mean by a physical metric is a metric tensor which properly encodes the notion of ‘distance between colors’ as perceived by an average human subject. The underlying (but often implicit) assumption of researchers in this field, going back at least to Helmholtz [7] and Schrödinger [8], appears to be that there is indeed a unique physical metric on  $\mathcal{C}$ . Whether this is indeed the case is not entirely clear to us. For, it may not be possible to make a *sufficient number* of experiments to pin down the metric tensor uniquely. Certainly, an average human subject does not have a *direct* perception that the distance between a pair of colors  $c_1$  and  $c_2$  is, say, twice the distance between another pair,  $c_2$  and  $c_3$ , even when the three colors are reasonably close to one another (so that issues such as ‘geodesic crossings’ do not arise). Therefore, if the metric can indeed be singled out uniquely, it would have to be via a series of a sufficient number of indirect experiments. What *is* clear is that experiments can be made to constrain the *form* of the possible metrics. In this section, we will provide an illustration. More precisely, we will see that there do exist experiments that show that  $\mathcal{C}$  is equipped with an additional structure, which in turn provides a handle on the *form* of allowable metrics. Note that this requires the use a new experimental input i.e., in addition to the form of the three adsorption coefficients used so far.

Recall that there is a well-defined vector field  $B^a$  on  $\mathcal{C}$ , the brightness vector field whose integral curves are the ‘radial lines’ in  $\mathcal{C}$ . The idea is to experimentally probe its properties which are tied to the physical metric on  $\mathcal{C}$ . More precisely, let us try to gain control on

some of the properties of the unknown metric by seeking directions which are *orthogonal* to  $B^a$ . Fix an arbitrary color  $c$  in  $\mathcal{C}$ . It lies on an integral curve  $\ell$  of  $B^a$ . Look at a ‘nearby’ integral curve  $\ell'$ , consider colors which lie on it and compare them to the original color  $c$ . Choose from all points on  $\ell'$  the point  $c'$  that seems ‘closest’ to  $c$ . Then, one can say that  $c$  and  $c'$  have the same brightness. Repeat this procedure for all integral curves that are close to  $\ell$ . When carried out by a large sample of human subjects with ‘normal’ vision, this experiment provides a two-dimensional ‘flat’ passing through  $c$ , *along which brightness is constant* for the average human subject. Along each integral curve of  $B^a$ , by contrast, it is only the brightness that changes. Thus, through each point  $c$  in  $\mathcal{C}$ , there is a two-flat which is *orthogonal* to the brightness vector field. Within experimental errors, these flats are ‘integrable’. That is, through each point  $c$  of  $\mathcal{C}$  there passes a two-dimensional surface which is everywhere orthogonal to the brightness vector field  $B^a$ . Thus we have a foliation of  $\mathcal{C}$  and hence we can introduce a *brightness function*  $B$  whose level surfaces are given by the foliation. A priori, the function  $B$  is not unique since we can replace it by any function  $f(B)$  of  $B$ . However, we can remove this arbitrariness by tying the function  $B$  with the vector field  $B^a$ . We will make the simplest choice  $B^a \partial_a B = 1$ , i.e., that the brightness function  $B$  is the affine parameter along the integral curves of the brightness vector field  $B^a$ .

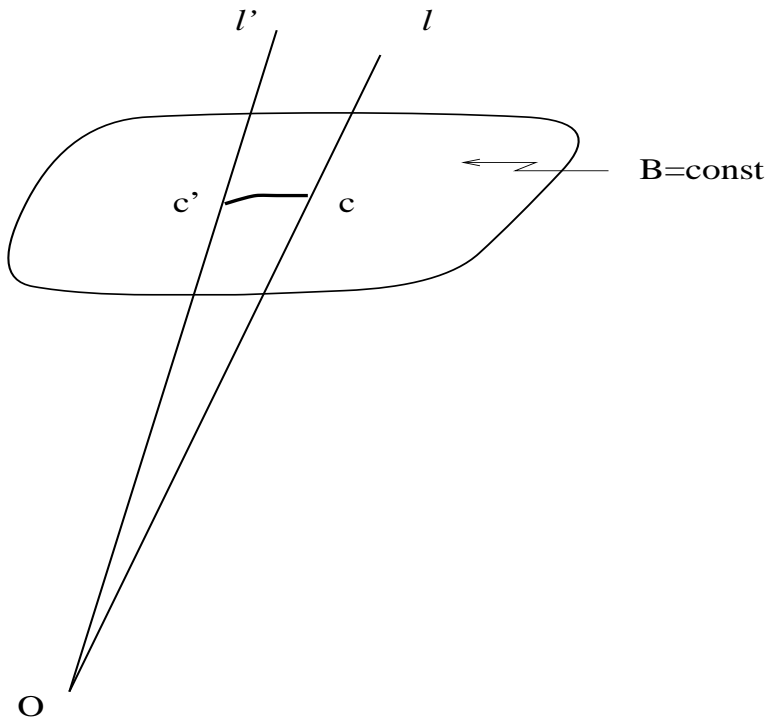


FIG. 3. The color  $c'$  is the point on  $\ell'$  that seems closest to the fixed  $c$ . The surface of constant brightness  $B = \text{const}$  is formed from such colors.

Let us restate the argument in a slightly different way. The manifold  $\mathcal{C}$  is equipped with a vector field  $B^a$ . The experiments indicate that the allowable metrics  $g_{ab}$  should be such that the co-vector field  $B_a := g_{ab} B^a$  is hypersurface orthogonal:  $B_{[a} \partial_b B_c] = 0$ . Note that the metric is essential in the formulation of this condition because the notion of hypersurface orthogonality refers to co-vector fields and *not* to vector fields. This condition is thus a restriction on the permissible metrics.

This new structure tells us that the metric  $g_{ab}$  has the following form:

$$g_{ab} = q_{ab} + N^2 \partial_a B \partial_b B \quad (2.10)$$

where  $q_{ab}$  is the intrinsic metric (or projection operator) on the two-manifolds of constant brightness and  $N^2 = g_{ab} B^a B^b$  is the squared-norm of the vector field  $B^a$ . Thus, while a general metric on  $\mathcal{C}$  has six arbitrary components, the allowable metrics have only four; three in  $q_{ab}$  and one in  $N$ . (In the relativity terminology, the 'shift vector' is zero.) Furthermore, since the constant brightness surfaces are two-dimensional and have trivial topology,  $q_{ab}$  is in fact conformally flat.

There *is* a mathematically natural metric on  $\mathcal{C}$  with respect to which  $B^a$  is hypersurface orthogonal as required. However, it is not clear to us that the brightness level surfaces determined by that metric are physically correct. Nonetheless, as an illustration, we will now sketch how this metric arises. This discussion will also bring out an interesting feature of the manifold  $\mathcal{C}$ .

In Sec. II B we provided the explicit expression (2.7) of the projection operator  $P$  from the space  $\mathcal{I}$  of intensities to the space  $\mathcal{C}$  of perceived colors. Let us consider  $\mathcal{I}$  as a convex-linear sub-space of the Hilbert space  $\mathcal{H} = L^2([-1, 1], d\lambda)$ , i.e., space of (real-valued) square-integrable functions on the interval  $(-1, 1)$ . This space is obviously equipped with a positive definite metric, the inner-products being given by:

$$\langle f_1, f_2 \rangle = \int_{-1}^1 d\lambda f_1(\lambda) f_2(\lambda). \quad (2.11)$$

The adsorption functions  $A_i(\lambda)$ , in particular, belong to  $\mathcal{H}$ . Denote by  $\mathcal{S}$  the three dimensional subspace of  $\mathcal{H}$  spanned by the  $A_i(\lambda)$ . The projection map (2.7) can be trivially extended to all of  $\mathcal{H}$ . For notational simplicity, we will denote the extension also by  $P$ . It is obvious from the explicit expression of this map that the image of  $\mathcal{H}$  under it is a three-dimensional vector space, which we will denote by  $V$ . It is straightforward to verify that the map  $P$  is an isomorphism between  $\mathcal{S}$  and  $V$ . Note that  $\mathcal{C}$  is a convex-linear sub-space of  $V$ . Denote by  $\mathcal{S}^+$  the convex-linear sub-space of  $\mathcal{S}$  obtained by projecting to  $\mathcal{S}$  the convex-linear space  $\mathcal{I}$  of all spectral distributions. (This is a Hilbert-space projection within  $\mathcal{H}$ .) Then, the image under  $P$  of  $\mathcal{S}^+$  is precisely  $\mathcal{C}$ . Since  $P$  is an isomorphism between  $\mathcal{S}$  and  $V$ , it is in particular an isomorphism between  $\mathcal{S}^+$  and  $\mathcal{C}$ . We will use this isomorphism to identify the two spaces. Thus, *not only is there a projection from  $\mathcal{I}$  to  $\mathcal{C}$ , there is also an embedding of  $\mathcal{C}$  (on to  $\mathcal{S}^+$  and therefore) in to  $\mathcal{H}$* . But since  $\mathcal{S}^+$  is a real, smooth sub-manifold of  $\mathcal{H}$ , it inherits from  $\mathcal{H}$  a positive definite, smooth metric. Hence  $\mathcal{C}$  is equipped with a metric  $g_{ab}^o$ . This is a natural metric at least from a mathematical point of view because the  $L^2$ -norm features also in the projection map (2.7).

Finally, let us discuss hypersurface orthogonality of  $B^a$ . The radial vector field is hypersurface orthogonal everywhere on  $\mathcal{H}$  (except at the origin), and hence also on any finite dimensional sub-space, such as  $\mathcal{S}$ , thereof. Restricted to  $\mathcal{S}^+$ , this is precisely the brightness vector field  $B^a$ . Thus, the positive-definite metric  $g_{ab}^o$  induced on  $\mathcal{C}$  from  $\mathcal{H}$  does have the required property that  $g_{ab}^o B^b$  is hypersurface orthogonal.

This particular metric  $g_{ab}^o$  on  $\mathcal{C}$  is flat. However, the natural coordinates,  $c_i$  do not provide a Cartesian chart for this flat metric because the adsorption functions  $A_i(\lambda)$  are

not mutually orthogonal in the  $L^2$ -norm. But the components of  $g_{ab}^o$  in the chart  $c_i$  can be written down readily once the  $L^2$ -inner products of the  $A_i(\lambda)$  are known. The latter can be estimated analytically rather easily by, for example, approximating the  $A_i(\lambda)$  by Gaussians. Indeed, in this approximation,  $\mathcal{S}$  can be thought of as the linear span of three quantum mechanical coherent states and calculations simplify considerably because of the available mathematical machinery associated with coherent states.

Finally, as explained in the Introduction, several other metrics have been proposed in the literature. However, they are invariably written in coordinates which are rather unnatural from the point of adsorption coefficients (and hence our simple model). It would take too long to explain the meaning of these coordinates here. Furthermore, none of the available metrics have a definitive status. Therefore, we will refrain from discussing them here. We wish to emphasize, however, that the problem of finding the complete description of the physical metric(s) on  $\mathcal{C}$  is extremely interesting from a practical viewpoint; its solution would be regarded as a major breakthrough with applications to a variety of industries including film, photography and color television.

### III. DISCUSSION

In mathematical physics, one generally begins by constructing simplified models in which the actual physical complications of the system under consideration are stripped down to a bare minimum—the oft-quoted example is that of a ‘spherical cow’. Such a strategy is generally essential because the model has to be sufficiently simple to be amenable to mathematical analysis. Furthermore, if these assumptions are physically well-motivated, the model generally leads to results which are correct at least qualitatively. Further improvements can be often made in increments by treating the ignored complications as perturbations.

We have proceeded in this spirit. Our basic assumption is that the essential features of perceived colors are all captured in the three adsorption coefficients  $A_i(\lambda)$ , whence the perceived colors are completely characterized by the three numbers  $c_i$  of (2.6). In the first approximation, this assumption is physically reasonable because  $c_i$  are the overlap integrals of the function  $I(\lambda)$  representing the incident energy flux per unit wavelength and functions  $A_i(\lambda)$  representing the fraction of the incident energy absorbed by the  $i$ -th cones. That is,  $c_i$  are a good measure of the signal that is sent to the brain.

However, as indicated already, our model fails to incorporate two important features: i) role of the rods in the retina which also contribute to the signal; and, more importantly, ii) signal processing in the brain itself. The role of rods is to enhance the shades of grey and they are therefore important in processing information that is received when the illumination is poor, e.g. in the evening and at night. In color perception which occurs in brightly illuminated environments, therefore, it seems safe to ignore their role. The role played by the brain, of course, is *much* more complex. However, these complexities seem not to be critical for ‘normal human subjects’ and under ‘normal viewing conditions’. That is, in this case, there appears to be a strong correlation between the final perception of the color and the input signal to the brain, coded in the three numbers  $c_i$ . This assertion is supported by the fact that our simple mathematical model reproduced all the general features of  $\mathcal{C}$  described in figure 1. Further support for the model comes from color blindness. Most color

blind people can see some colors but not all. In terms of the model, they have only two types of cones rather than all three. Their space of perceived colors is thus only two dimensional. A very small fraction perceive no colors at all. They have only one type of cones and so their space of colors is only one dimensional. All this is in agreement with observations.

Nonetheless, it is important to remember that color perception is in fact a ‘psycho-physical’ phenomenon while the model treats it simply as a physical phenomenon. Thus, it is quite incomplete. What does it miss? Here is a simple example. Let us consider an apple. Under normal illuminations –e.g., in reasonably bright sunlight– the model works well. Under these ‘normal’ conditions, the spectral distribution  $I(\lambda)$  reflected by the apple is highly peaked around red, whence, according to the model only  $c_1$  is significant and we do indeed perceive it as being red. However, the apple can be illuminated by artificial light and then the reflected intensity  $I(\lambda)$  can be quite different. So, our model would now predict that all three  $c_i$  should be non-zero and hence we should perceive quite a different color. This does *not* happen for a large set of illuminations; an average human subject continues to perceive a red apple! This is because the brain now performs further calculations to re-interpret the incoming signal. Such re-interpretations are essential for our survival which requires that we should be capable of recognizing the ‘sameness’ of the object even when it is illuminated quite differently. There are dozens of day to day examples in which the brain interferes non-trivially to make the model completely inadequate. Snow, in particular, looks white even when we are walking in a pine forest, although now it in fact reflects predominantly green light. Similarly, for a large class of illuminations, a patch color can seem to be green in certain surroundings but pale grey when looked at in isolation. A sheet of paper appears white in both bright day light and in a windowless office even though the amount of light reflected from a dark grey sheet in daylight is much more than that reflected by the ‘white’ sheet in a windowless office. In each case, the brain has absorbed not just the three inputs  $c_i$  but a whole range of other ‘relevant’ information to produce a final color perception. By now, neurologists even know the place in the cortex which responds to color as opposed to wavelengths. It is a pea-sized region, called  $V_4$ , which is distinct from another region  $V_1$  that responds to specific wave-lengths, or, in terms of our model, to the three  $c_i$ .

A striking example of this inadequacy of our purely physical model was seen recently. It is described by Oliver Sacks in his poignant essay *The case of the colorblind painter* [10]. This is a story of a successful artist who had been painting for over 40 years. In 1986, he had a minor car accident. He suffered a concussion, but recovered rather quickly except for one thing: he lost his color vision! Detailed examination showed that his retina –with all its three types of cones– were in tact and functioning normally. So, his brain was receiving the color signals as before; the space of colors *received* in the brain continued to be three-dimensional. However, the space of *actually perceived* colors was now only one dimensional! Examination showed that the whole brunt of damage to the cortex from the accident was borne precisely by the  $V_4$  area.

To summarize, our simple model works well when we study colors in isolation, e.g., when a uniform field of color is presented to the average eye on a black background. In the ‘real world’ however factors such as illumination, the state of adaptation of the visual system and neighboring stimuli in the visual field cause the brain to intervene and remove correlations between the input signal (the three  $c_i$ ) and the actual color perception. These interventions are incredibly fascinating processes. But they also seem to be among the most complex

ones. It will probably take a long time before such processes can be successfully coded in the mathematical models of color perception.

There is another fascinating problem that has attracted considerable attention: the role of color perception in the evolution of species. One might have naively thought that color perception becomes more and more sophisticated as the species become more advanced in evolutionary terms. Surprisingly, this appears *not* to be the case! The space of perceived colors of a frog, for example, is *four* dimensional—their eyes have four types of cones. Could it be that there is actually an anti-correlation between the number of colors perceived and place of the species in the evolutionary ladder? Again, the answer is in the negative! While the dimension is four for frogs and three for humans, it is zero for most mammals. Thus, most mammals, including the big cats which are such excellent hunters, are color blind. Naively, one might have thought that, because they have to hunt in diverse environments, evolutionary forces would have endowed them with a large dimensional color space. This is simply not the case. While there do exist interesting attempts to understand the role of evolution and natural selection in color perception (see, e.g., [11]), this is still a largely open issue.

**Acknowledgments** This work was supported in part by the NSF grant PHY 95-14240, by the Eberly Research funds of The Pennsylvania State University and by DGAPA of the Universidad Nacional Autónoma de México

## REFERENCES

- [1] I. Newton, Phil. Trans. Roy. Soc. London **80**, 3075 (1671); I. Newton, *Opticks*, Smith and Walford, London (1704).
- [2] T. Young, Phil. Trans. Roy. Soc. London **92**, 20 (1802).
- [3] D. Mitchell and W. Rushton, Vision Research **11**, 1033 (1971).
- [4] H. Grassmann, Ann. Phys. Chem. (peggendorf) **89**, 69 (1853); translation in [5].
- [5] D.L. MacAdam, *Sources of Color Science*, MIT Press, Cambridge Mass (1970).
- [6] B. Riemann, *Ueber die Hypothesen, Welche der Geometrie Zu Grunde liegen* Gött. Nachr. **13**, 133 (1868).
- [7] H. Helmholtz, Sitzungber. d. Berl. Akad. Dec 17, 1071 (1891).
- [8] E. Schrödinger, Annalen del Physik **63**, 397 (1920); *ibid*, 481 (1920).
- [9] G. Wyszecki and W.S. Stiles, *Color Science*, Wiley, New York (1982).
- [10] O. Sachs, *An Anthropologist on Mars*, (Vintage Books, New York, 1996) pages 3-41.
- [11] J.W. Weinberg, Gen. Rel. Grav. **7**, 135 (1976).