The Theory of the Maximum Visual Efficiency of Colored Materials

DAVID L. MACADAM, *Massachusetts Institute of Technology*

(Received May 20, 1935)

The term "visual efficiency" is used here to represent interchangeably the concepts of visual transmission factors of filters and of visual reflection factors of reflecting surfaces. Fluorescent substances are excluded from consideration. A material exhibits hue because visible radiation in some wavelength bands is more or less completely absorbed. This partial absorption of incident energy, necessary for the appearance of hue, obviously decreases the visual efficiency below the unit efficiency characteristic of a nonabsorbing, hueless, white material. Of all the conceivable spectrophotometric curves of materials exhibiting a given chromaticity when illuminated with light of a specified quality, there must be at least one which yields a maximum value for the visual efficiency. This paper describes the general type of spectrophotometric curve which is known to have this unique property. A new proof of the validity and uniqueness of this type of curve is presented. This proof takes advantage of the simplifications made available by the adoption of the I. C. I. 1931 coordinate system for colorimetry.

FROM his experimental survey of the chromaticity range of real pigments, Ostwald[1] drew the empirical conclusion that the attainment of maximum saturation with a pigment required that the spectral reflectance have only the values unity or zero (with discontinuous transitions of the spectrophotometric curve between these values). Furthermore, he concluded that such pigments should have either a single continuous transmission band or a single continuous absorption band in the visible region of the spectrum.

Schrödinger[2] examined the problem of specifying the spectrophotometric characteristics of pigments which should have the maximum attainable reflectance for a given chromaticity (dominant wavelength and purity). His discussion was based on the fundamental properties of the normal human observer as revealed by color mixture experiments. From the convexity of the spectral locus specified by these color mixture data for any coordinate system, and from the additive properties of color mixture diagrams so constructed, he succeeded in establishing a logical confirmation of Ostwald's empirical rule.

Luther,[3] Nyberg[4] and Rösch[5] have computed the loci in the color mixture diagram which bound all chromaticities attainable with pigments having specified visual reflection factors under various illuminants. Independently of these workers, Gage[6] computed the maximum attainable transmission of highly saturated red filters and presented the methods and results before a meeting of the Optical Society of America. All of these computations were made prior to the adoption of the I. C. I. coordinate system for colorimetry.[7]

The conclusions of Ostwald and of Schrödinger can be regarded as corollaries of the following theorem: The maximum attainable purity (closest approach to the corresponding spectral color) for a material having a specified dominant wavelength and visual efficiency will be attained if the material has a spectrophotometric curve which is everywhere either zero or unity, and which has at most two transitions between these values within the region of visible radiation. This theorem is valid for any definition which causes purity to increase continuously during the transition from the neutral toward fully saturated spectral or purple hues. In particular, excitation purity which is defined exclusively by and for a specified coordinate system as the fraction of the distance from the neutral toward the associated saturated hue, spectral or purple, satisfies this condition.

---

[1] W. Ostwald, Königle Sächs. Ges. d. Wiss. Abh. d. Math.-Phys. 34, 471 (1917); Physik. Zeits. 17, 328 (1916).

[2] Erwin Schrödinger, *Theorie der Pigmente von grösster Leuchtkraft*, Ann. d. Physik 62, 603 (1920).

[3] Luther, Zeits. f. tech. Physik 8, 540 (1927).

[4] Nyberg, Zeits. f. Physik 52, 407 (1928).

[5] S. Rösch, Fortshritte der Mineral., Krist. & Petr. 13, 143 (1929).

[6] H. P. Gage, J. O. S. A. and R. S. I. 18, 167 (1930).

[7] D. B. Judd, *The I. C. I. Standard Observer and Coordinate System for Colorimetry*, J. Opt. Soc. Am. 23, 359 (1933).

The proof of this theorem assumes a particularly simple and readily comprehended form if the problem is shown to be equivalent to a certain mechanical problem. The equivalence of the colorimetric problem to this mechanical problem makes it possible to use familiar theorems concerning centers of gravity and moments in order to deal expeditiously with the relatively unfamiliar color problem. The notation used follows that adopted by Judd with a few slight modifications and additions.

$\bar{x}_\lambda$, $\bar{y}_\lambda$, $\bar{z}_\lambda$ are the I. C. I. color mixture functions giving the amounts of three unitary stimuli necessary to produce, in an additive mixture, a visual color match with a unit amount of radiation of wavelength λ.

E_λ is the coefficient of $d\lambda$ giving the power (per unit area) of radiation incident on the sample and having wavelengths in the region from λ to $\lambda+d\lambda$.

r_λ is the ratio of the power of the radiation of wavelength λ reflected or transmitted by the sample to the power of the radiation of the same wavelength incident on the sample.

X, Y, Z are the amounts of the above-mentioned unitary stimuli necessary to match in color the light reflected or transmitted by the sample.

$$X=\int_0^\infty \bar{x}_\lambda E_\lambda r_\lambda d\lambda,\ Y=\int_0^\infty \bar{y}_\lambda E_\lambda r_\lambda d\lambda,\ Z=\int_0^\infty \bar{z}_\lambda E_\lambda r_\lambda d\lambda.$$

x_λ, y_λ are the coordinates in the color mixture diagram of the point representing the chromaticity of radiation of wavelength λ.

$$x_\lambda=\bar{x}_\lambda/(\bar{x}_\lambda+\bar{y}_\lambda+\bar{z}_\lambda),\ y_\lambda=\bar{y}_\lambda/(\bar{x}_\lambda+\bar{y}_\lambda+\bar{z}_\lambda).$$

x, y are the coordinates in the color mixture diagram of the point representing the chromaticity of the sample when illuminated with light having the energy distribution E_λ. By analogy with x_λ, y_λ and consistent with the additive properties of vision,

$$x=X/(X+Y+Z),\quad y=Y/(X+Y+Z).$$

From the definitions:

$$(X+Y+Z)=\int_0^\infty(\bar{x}+\bar{y}+\bar{z})_\lambda E_\lambda r_\lambda d\lambda=\int_0^\infty m_\lambda d\lambda.$$

The symbol $m_\lambda=(\bar{x}+\bar{y}+\bar{z})_\lambda E_\lambda r_\lambda$ has been introduced as an abbreviation in this expression. From the expressions for x_λ, y_λ it is evident that

$$\bar{x}_\lambda E_\lambda r_\lambda=x_\lambda m_\lambda,\quad \bar{y}_\lambda E_\lambda r_\lambda=y_\lambda m_\lambda.$$

Consequently.

$$X=\int_0^\infty x_\lambda m_\lambda d\lambda \quad\text{and}\quad Y=\int_0^\infty y_\lambda m_\lambda d\lambda,$$

and

$$x=\int_0^\infty x_\lambda m_\lambda d\lambda\Big/\int_0^\infty m_\lambda d\lambda,$$

$$y=\int_0^\infty y_\lambda m_\lambda d\lambda\Big/\int_0^\infty m_\lambda d_\lambda.$$

If m_λ is now regarded as a mass per unit wavelength interval along the spectrum locus in the color mixture diagram, x, y are evidently the coordinates of the center of gravity of this mass distribution, defined by

$$m_\lambda=(\bar{x}+\bar{y}+\bar{z})_\lambda E_\lambda r_\lambda.$$

For the normal observer and a given illuminant, m_λ has a maximum value corresponding to $r_\lambda=1$ and a minimum value, zero. Under the same conditions the visual efficiency is proportional to Y because the visibility function is identical with $\bar{y}$. Consequently visual efficiency, R, is proportional to the moment of the mass distribution about the x axis, $R=\text{const.}\int_0^\infty y_\lambda m_\lambda d\lambda$.

Consider a material having the spectrophotometric curve indicated by the full line in Fig. 1a. The point representing the chromaticity of this material is the center of gravity of the spectrum locus weighted in the above manner to the greatest possible extent everywhere between λ_1 and λ_2; ($r_\lambda=1$). Let this center of gravity be C_1. The dominant wavelength, λ_0, is the wavelength associated with the point on the spectrum locus collinear with C_1 and the white point, see Fig. 2. The white point is the center of gravity of the spectrum locus weighted everywhere to the maximum extent possible for the illuminant considered. The visual efficiency is $R=k\int_0^\infty y_\lambda m_\lambda d\lambda$. The dotted line in Fig. 1 indicates a modification of this spectrophotometric curve of such a character as to maintain λ_0 and

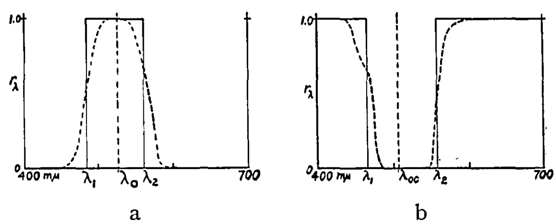

FIG. 1a. Spectrophotometric curves for two green samples having dominant wavelength about 525 mμ and visual efficiency about 0.50. Box shaped curve, full line, has maximum attainable purity.

FIG. 1b. Spectrophotometric curves for two purple samples complementary to 505 mμ and having visual efficiency about 0.50.

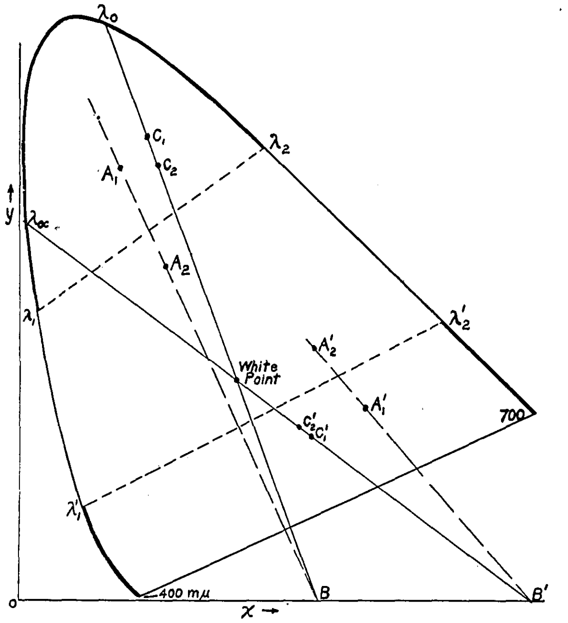

FIG. 2. Color mixture diagram according to I.C.I. standardization, with construction used in proof that box shaped spectrophotometric curve results in maximum purity for every specified dominant wavelength and visual efficiency.

R constant. C_2 is the center of gravity of the distribution determined by this altered spectrophotometric data. Subtraction of mass from the original distribution is the only conceivable change that can be made in the region for which r_λ was originally unity. Suitable masses must be added outside of the limits of the original distribution in order to keep R unchanged. The amounts of these added masses depend on their distribution and on the amounts and distribution of the subtracted masses. These added masses are determined by the condition that λ_0 and R remain constant. The center of gravity of the subtracted masses is at A_1. The point A_2 is the center of gravity of the added masses. All of the subtracted masses, and consequently their center of gravity A_1, must lie on the same side of the line connecting λ_1 and λ_2 as does C_1. Similarly, A_2 must lie on the opposite side of this line. The point B is the center of gravity of A_1 and A_2 weighted, respectively, with $-m_1$ (the total subtracted mass) and $+m_2$ (the total added mass). The point B must lie along the λ_0 locus because C_2, the center of gravity of C_1 and B, must lie on the λ_0 locus. The point B must also lie on the x axis because the moment of m_2-m_1 about the axis must be zero in order to keep R constant. This same condition requires that m_2-m_1 be positive when A_1 is farther from the x axis than A_2, and negative when A_2 is farther from the x axis than A_1. It is therefore evident from the diagram that if C_1 is farther than the white point from the x axis, m_2-m_1 will be positive. Since the total mass has increased, while the moment of this mass about the x axis has remained fixed, C_2 must lie closer to the x axis than does C_1. Consequently C_2 lies closer to the white point than does C_1. If C_1 is lower than the white point, m_2-m_1 will be negative and C_2 again will be closer to the white point than C_1. In all cases the variation of the spectrophotometric curve from the original (solid curve) results in a decrease in purity for an assigned λ_0 and R. The same argument will evidently apply for the purple material whose spectrophotometric curve is shown in Fig. 1b. Fig. 2 contains the construction necessary for following the above reasoning in this case. The primed letters refer to the purple case. If purity is measured so as to increase continuously as C_1 becomes more distant from the white point, along any straight line through the white point, then the terminology for spectral colors will carry over to the discussion of the purples.

The above reasoning has shown that there is a maximum attainable purity corresponding to every specified pair of values of dominant wavelength and visual efficiency. This maximum purity is attained only if the spectrophotometric data of the colored material are either zero or unity at every wavelength in the visible spectrum, with no more than two transitions between these values within the visible range. Of two such samples having the same dominant wavelength, the one exhibiting the higher efficiency invariably exhibits the lower purity. Consequently the fact that there is a maximum attainable efficiency determined by each specified chromaticity can be regarded as a corollary of the maximum purity theorem. The spectrophotometric characteristics necessary to attain this maximum efficiency are identical with those necessary to attain maximum purity. Maximum efficiency and maximum purity can be regarded interchangeably as functions of each other, the nature of the functions depending on dominant wavelength.

A subsequent paper will present the data for the dependence of maximum visual efficiency on chromaticity. These data will be represented in tabular form with dominant wavelength and excitation purity as chromaticity variables, and also in charts showing the loci of chromaticities characterized by equal maximum efficiencies traced in the I. C. I. color mixture diagram. These tables and charts will be given for I. C. I. illuminants *A* and *C*.

The author wishes to express his appreciation of the interest and assistance of Professor Arthur C. Hardy during the preparation of this paper.