Four-Transition 0-1 Functions for Reflectance Spectra

Paul Centore

© February 25, 2018

Abstract

It is known that a reflectance spectrum for an optimal colour takes on the value 0 or 1 at every wavelength, with at most two transitions between those values. This paper shows that any non-optimal colour can be produced by a reflectance spectrum that takes on the value 0 or 1 at every wavelength, with at most four transitions. While the two-transition optimal spectrum is unique, the four-transition non-optimal spectrum is not unique.

1 Introduction

Colorimetry standardizes the results of a colour-matching experiment, in which a viewer observes, through an aperture, two colour stimuli against a black background. Typically the stimuli occupy two semicircles which meet along their diameter. In these experiments, two stimuli with very different spectral distributions can produce identical colour sensations, making the semicircles' colours match. In 1931, the Commission Internationale de l'Éclairage (CIE) introduced[1] a three-dimensional XYZ colour space, and a set of colour-matching functions which assign a vector in that space to a physical stimulus. Two colour stimuli in an aperture experiment match if and only if their XYZ coordinates are identical. In colorimetry, then, although not in more realistic viewing situations where chromatic adaptation comes into play, one can say that a stimulus has a unique colour given by XYZ. This paper works in the context of colorimetry, and so considers colours in that sense.

For any illuminant I, the terms defined in colorimetry allow the construction of an object-colour solid[2] Ω_I, which is a subset of XYZ space. The object-colour solid for I consists of all the XYZ coordinates that can be produced when a human views a matte surface that is illuminated by I. A surface has a *reflectance spectrum*, which is a function that specifies, for each visible wavelength (roughly those between 400 and 700 nm) the percentage of incoming light of that wavelength that the surface reflects. A reflectance spectrum takes on values between 0 and 100%, or, more simply, 0 and 1. Typically, many different spectra, making up a *metamer set*, can produce the same XYZ when illuminated by I.

The colours on the boundary of Ω_I are called *optimal colours*. A well-known result,[3,4] referred to here as the Optimal Colour Theorem, states that an optimal colour is produced only by a reflectance spectrum in *Schrödinger form*: the spectrum takes on only the values 0 and 1, and there are at most two transitions from 0 to 1 or vice versa. For convenience, a function that takes on only the values 0 or 1 will be said to have *0-1 form*, and the

standard term *support* will denote the set of wavelengths where a spectrum has non-zero reflectance. To provide a further simplification, a three-dimensional geometric interpretation suggests joining the ends of the spectrum to produce a circular domain of wavelengths; on the circular domain, the support of an optimal spectrum is an interval, and, excepting ideal black and ideal white, each optimal spectrum has exactly two transitions. The theorem includes a converse: any spectrum in Schrödinger form must produce an optimal colour. Further work[5-8] has shown that the reflectance spectrum for an optimal colour is unique.

A natural question is whether any similar simple forms occur for non-optimal colours, those that are inside the object-colour solid. This situation is more complicated because any non-optimal colour can be produced by a large *metamer set* of varied reflectance spectra. This paper will show, however, that any non-optimal colour can be produced by a 0-1 reflectance spectrum with *four* transitions, as opposed to the two that characterize optimal colours.

Four-transition spectra can arise naturally by adding optimal spectra. Two optimal spectra whose supports do not intersect will be called *disjoint*. Adding two disjoint optimal spectra produces a 0-1 function whose support consists of two intervals. This spectrum has four transitions, one at the endpoint of each interval. Given a non-optimal colour, in the interior of an object-colour solid, we will find two disjoint optimal colours, on the boundary of the solid, whose vectors in XYZ space sum to the vector of the non-optimal colour in that space. It then follows that the sum of the reflectance spectra, which is a four-transition 0-1 function, must be metameric to the non-optimal colour, proving the result.

The selection of the pair of disjoint optimal colours, which is not unique, uses the convexity and central symmetry of object-colour solids.[6] Any section of the solid that contains the center, the origin, the white point, and the non-optimal colour, is also convex and centrally symmetric. The colours on the boundary of the section must be optimal colours on the original object-colour solid, and the non-optimal colour is interior to the section. The symmetry will allow pairs of disjoint optimal colours to be characterized geometrically, and a construction involving translations will show that any vector in the interior is the sum of two vectors, given by disjoint colours, on the boundary.

The paper is organized as follows. After the introduction, the perception of object or surface colours is formulated geometrically. Object-colour solids are of particular interest; their convex, centrally symmetric strucutre, and the inheritance of that structure by central sections of the solid, are described. Reflectance spectra are expressed in a circular form, which simplifies the counting of transitions, and implies that any optimal spectrum (excepting black and white) is a 0-1 function with exactly two transitions. An expression of a four-transition 0-1 function as the sum of two disjoint two-transition 0-1 functions is given. The geometric constructions involving a section of an object-colour solid are then used to write an arbitrary non-optimal colour as the vectorial sum of two disjoint optimal colours in XYZ space, from which the main result of the paper follows: the metamer set of an arbitrary non-optimal colour contains a four-transition 0-1 reflectance spectrum. The paper closes with a summary.

2 Geometric Formulation

2.1 Reflectance Spectra

An object's colour is perceived when light reflects off that object and enters a viewer's eye. The light must be in the visible spectrum, and can be described by a spectral power distribution (SPD), which is a density function over the visible spectrum; integrating this function over any interval gives the power of the light in that interval. The boundaries of the visible spectrum are not hard and fast, and this paper will use 400 and 700 nm. Other sources use different values, such as 360 and 760 nm; the development in this paper can readily be extended to accommodate different endpoints. In addition, for consistency with colorimetry, reflection will be assumed to be matte, also known as diffuse or Lambertian, rather than specular. An *illuminant* is an SPD that gives the relative power of a light source at various wavelengths. Illuminants are used rather than absolute SPDs, because the intensity of a light typically varies over a scene, while its SPD's shape remains the same.

When light strikes an object, some of that light is reflected, and the rest is absorbed or transmitted, the proportions of each outcome varying with wavelength. An object's *reflectance spectrum* gives the percentage of light that that object reflects, as a function of the visible wavelengths. A spectrum is bounded between 0 and 100%, or equivalently, between 0 and 1. An object modifies the SPD of incoming light, by reflection in accordance with its spectrum, to produce a new SPD, which arrives at an observer.

The CIE 1931 Standard Observer[1] systematizes human perception of an incoming SPD. This standard gives three colour-matching functions ($\bar{x}(\lambda)$, $\bar{y}(\lambda)$, and $\bar{y}(\lambda)$), each of which is linear on the set of SPDs. The function's outputs are three non-negative CIE coordinates, denoted X, Y, and Z, which can be seen as a vector in $\mathbb{R}^3$. A viewer in a colour-matching experiment will perceive that two SPDs have the same colour if and only if their XYZ coordinates are equal. (In more realistic settings, chromatic adaptation might cause perceptual differences, but this paper is restricted to the setting of standard colorimetry.) Two reflectance spectra that produce the same colour perception are called *metamers*, and the set of all reflectance spectra that produce the same XYZ when viewed under a certain illuminant I are called the *metamer set* of XYZ for I.

2.2 Object-Colour Solids

Not all XYZs are possible, because there might not exist *any* SPD which produces a particular XYZ. Furthermore, not all XYZs involve an object: an observer can also view a light source directly, without any intermediate reflection. The set of all possible XYZs that do result when an illuminant I strikes an arbitrary matte object is called the *object-colour solid* Ω_I for I, and is a subset of $\mathbb{R}^3$.

Even though different illuminants lead to different solids, all object-colour solids share some geometric features. For instance, they all start at the origin, which corresponds to the reflectance spectrum that is identically 0; such an *ideal black* object would reflect no light at all. At its brightest, on the other hand, each solid terminates at an *ideal white* point, corresponding to the reflectance spectrum that is identically 1. The origin and white point are extreme vertices of the solid. By linearity, the line segment joining them, called the

neutral axis, is also part of the solid.

Every object-colour solid, furthermore, is convex, and centrally symmetric about $XYZ_{0.5}$, the vector that results from the reflectance spectrum that is 0.5 at every wavelength. A set $\mathcal{S}$ is *convex* if the line segment between any two points in $\mathcal{S}$ is also contained in $\mathcal{S}$. Convex sets therefore admit no bumps, indentations, holes, or internal cavities. By linearity, $XYZ_{0.5}$ lies on the neutral axis, halfway between the origin and the white point. Since convexity and central symmetry will be needed later, we will demonstrate them directly.

To show that Ω_I is convex, choose any two of its points; denote them $XYZ_1 = (X_1, Y_1, Z_1)$ and $XYZ_2 = (X_2, Y_2, Z_2)$. Then a point on the line segment between them is a vector $XYZ_\alpha = (X_\alpha, Y_\alpha, Z_\alpha)$ of the form

$$(X_\alpha, Y_\alpha, Z_\alpha) = \alpha(X_1, Y_1, Z_1) + (1 - \alpha)(X_2, Y_2, Z_2), \tag{1}$$

where α is between 0 and 1. Since XYZ_1 and XYZ_2 are both in Ω_I, they must result from reflectance spectra ρ_1 and ρ_2. Now construct a new reflectance spectrum

$$\rho_\alpha = \alpha\rho_1 + (1 - \alpha)\rho_2. \tag{2}$$

Since X, Y, and Z are linear functions of reflectance spectra, it follows that

$$X(\rho_\alpha) = \alpha X(\rho_1) + (1 - \alpha)X(\rho_2), \tag{3}$$
$$Y(\rho_\alpha) = \alpha Y(\rho_1) + (1 - \alpha)Y(\rho_2), \tag{4}$$
$$Z(\rho_\alpha) = \alpha Z(\rho_1) + (1 - \alpha)Z(\rho_2), \tag{5}$$

so the XYZ coordinates for ρ_α are just as given in Equation (1). Since these coordinates result from the reflectance spectrum ρ_α, they must occur in the object-colour solid. This argument applies to any α, so the entire line segment is contained in the object-colour solid, proving convexity.

Central symmetry, about the coordinates $XYZ_{0.5}$ of the reflectance spectrum

$$\rho_{0.5}(\lambda) = 0.5, \tag{6}$$

whose value is uniformly 0.5 at every visible wavelength λ, can also be shown directly. Consider an arbitrary point XYZ_1 in an object-colour solid. Then it resulted from some reflectance spectrum ρ_1. Now define the reflectance spectrum

$$\rho_2(\lambda) = 1 - \rho_1(\lambda). \tag{7}$$

Then

$$\rho_2(\lambda) + \rho_1(\lambda) = 1 \tag{8}$$
$$0.5(\rho_2(\lambda) + \rho_1(\lambda)) = 0.5, \tag{9}$$

so

$$\rho_{0.5}(\lambda) = 0.5(\rho_2(\lambda) + \rho_1(\lambda)). \tag{10}$$

After applying the linear transformations X, Y, and Z, to both sides of Equation (10), we find that $XYZ_{0.5}$ is exactly halfway between XYZ_1 and XYZ_2. Therefore the point XYZ_2 is diametrically opposite XYZ_1, relative to $XYZ_{0.5}$. Since the point XYZ_1 was arbitrary, every point of the object-colour solid has such a diametrically opposite, or *complementary* point, so an object-colour solid is centrally symmetric, as was to be shown.

2.3 Optimal Colours

As mentioned earlier, a particular XYZ can be produced by a metamer set of reflectance spectra. Since the set of reflectance spectra has many degrees of freedom, while CIE XYZ space only has three, metamer sets are often very large: a surprisingly wide variety of very different spectra can produce the same XYZ. An important exception to this case occurs for the colours on the boundary of an object-colour solid, which are known as *optimal colours*. The Optimal Colour Theorem, which Schrödinger[3] stated and proved in 1920, and MacAdam[4] derived by a physics-based approach in 1935, states that the reflectance spectrum for an optimal colour must have *Schrödinger form*: it takes on only the values 0 and 1, with at most two transitions between those values.

Schrödinger[3,10] classified the optimal colour spectra into the four forms shown in Figure 1, depending on the number and locations of transition wavelengths. The first form, in the upper left, has one transition, and takes on the value 1 between 400 nm and that transition. The second form, in the upper right, also has one transition, but takes on the value 1 to the right of the transition. The third form, in the lower left, has two transitions, between which it is 1, and outside which it is 0. The fourth form also has two transitions, but is 0 between them, and 1 outside.

These four forms can be conveniently subsumed into one form,[9] by joining the ends of the visible spectrum, at 400 and 700 nm, to form a circle. On the circle, the reflectance function for any optimal colour fills out a sector of an annulus, with one transition wavelength where the sector starts, and another where the sector ends. In the first form, the sector starts at 400 nm (which is also 700 nm). In the second form, the sector ends at 400 (or 700) nm. In the third form, the sector does not cross the 400 nm mark. In the fourth form, the sector starts before 700 nm, continues clockwise across 700 nm, which is also 400 nm, and then fills out some of the clockwise side of the 400 nm mark. An optimal reflectance spectrum takes on value 1 at wavelengths whose spectrum locus vectors are summed up to produce that optimal colour, and the Optimal Colour Theorem says that those wavelengths must lie along an interval of the circular spectrum locus.

In the circular form, any of the four cases in Figure 1 results in exactly two transitions. Equivalently, we could count an extra transition in Figure 1 whenever the spectrum is 0 at the left end and 1 at the right end, or vice versa. With this understanding, it is impossible for a spectrum to have exactly one transition, or indeed any odd number of transitions. Only the spectra for ideal white and ideal black have zero transitions. If we leave aside these two special cases, then the Optimal Colour Theorem says that a reflectance spectrum for an optimal colour must be a 0-1 function with exactly two transitions.

The converse of the Optimal Colour Theorem also holds: any reflectance spectrum in Schrödinger form produces an optimal XYZ, on the boundary of Ω_I. Further work has strengthened this theorem to include uniqueness: exactly *one* reflectance spectrum produces a particular optimal colour. (The history of the proof of uniqueness is unclear. In a footnote to a 2009 paper,[7] A. Logvinenko refers to a proof in an unpublished manuscript[8] that was still under preparation. An explicit proof of uniqueness,[5] which relies on some empirical findings of linear independence, was given in 2012; an extended discussion of the 2012 derivation appears in Ref. 6.) Uniqueness implies that the metamer set of an optimal colour consists of a single reflectance spectrum.
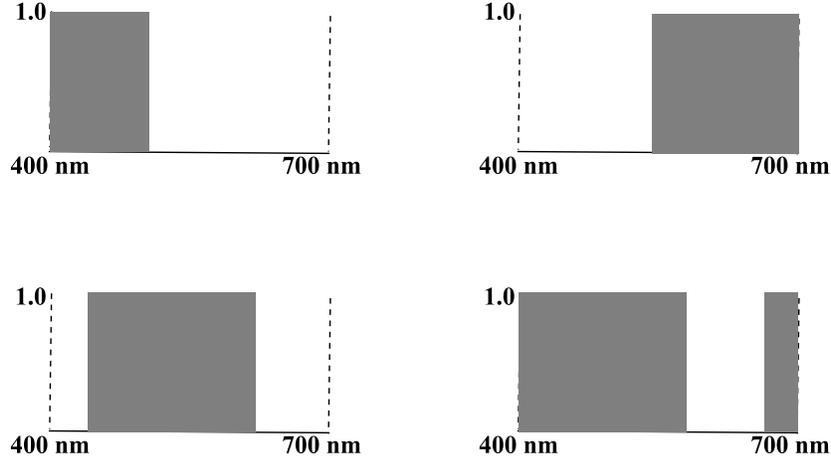
Figure 1: Four Forms for Optimal Colour Functions
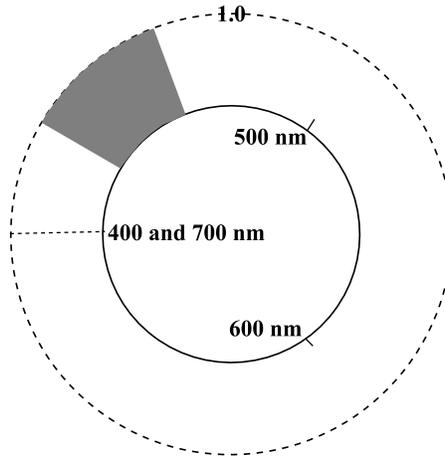
Figure 2: Circular Form for Optimal Colour Functions

A natural question is whether similar characterizations exist for non-optimal colours, those that are interior to Ω_I. This paper will provide such a characterization: every metamer set for a non-optimal colour contains at least one 0-1 reflectance spectrum with exactly four transitions. Unlike in the optimal case, however, the four-transition 0-1 spectrum is not unique, and, of course, many more metamers take on values besides 0 and 1.

2.4   Sections of an Object-Colour Solid

The derivation of the four-transition 0-1 result will consider not just object-colour solids, but also sections of those solids, in particular the sections given by intersecting the solid with a plane that contains both the origin and the white point. Since such a plane contains the central point, the section must be centrally symmetric. Since the solid and the cutting plane are both convex, the section is also convex. Figure 3 shows an example, with the origin labeled 0 and the white point labeled w. The center is also indicated. For convenience, the boundary of the section can be divided into upper and lower parts.
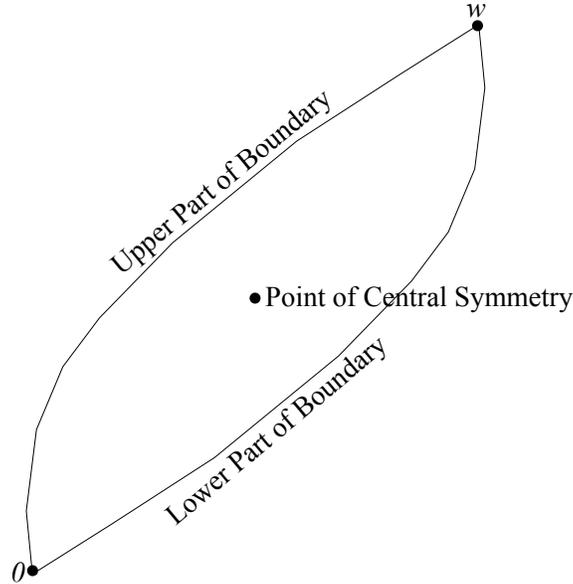
Figure 3: A Section of an Object-Colour Solid, Containing the Origin and White Point

The points on the boundary of the section are also on the boundary of the object-colour solid, so the Optimal Colour Theorem says that they have unique reflectance spectra in Schrödinger form. At the origin, of course, the spectrum is identically 0 (i.e. its support is empty), while at the white point it is identically 1 (i.e. its support consists of all visible wavelengths). As one moves along either part of the boundary from 0 to w, the support must always be an interval (allowing for wraparound), and it expands smoothly, sometimes to the left and sometimes to the right and sometimes to both left and right simultaneously, until it fills the spectrum from 400 to 700 nm. The support at any one point is always a proper subset of the support at a later point on that part of the boundary.

Though the starting and ending intervals for the upper and lower boundaries are the same, the intermediate intervals can be very different. In fact, the symmetry implies a complementary relationship. Let A and B be two diametrically opposite boundary points on the upper and lower parts, as shown in Figure 4. Since they are symmetric about the center, their sum must be identically 1, which is the spectrum for ideal white:

$$\rho_w = \rho_A + \rho_B. \tag{11}$$

The reflectance spectra for A and B are complements: ρ_A is 1 wherever ρ_B is 0, and 0 wherever ρ_B is 1. A corollary is that ρ_B is *disjoint* from any point on the boundary between 0 and A, where two spectra are said to be disjoint if their supports have empty intersection.

3   Derivation of Main Result

This section will derive the main result of this paper: the metamer set of any non-optimal colour contains a four-transition 0-1 reflectance spectrum. First, a four-transition 0-1 function will be constructed as the sum of a pair of disjoint Schrödinger functions. Then, such
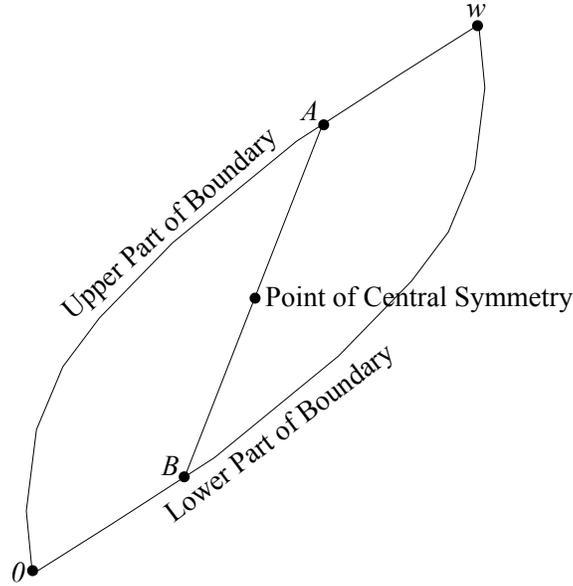
Figure 4: Two Diametrically Opposite Points on a Section of an Object-Colour Solid

a pair will be found geometrically, in a section of the object-colour solid, that sum to an arbitrary non-optimal colour. The result follows.

3.1   Constructing Four-Transition 0-1 Functions

Suppose we have two disjoint optimal colours, γ_1 and γ_2. Then the supports of their reflectance spectra ρ_1 and ρ_2 are two non-intersecting intervals, and their sum is a third spectrum:

$$\rho = \rho_1 + \rho_2. \tag{12}$$

Note that the sum of two reflectance spectra is not generally another reflectance spectrum, because the sum could result in values that are greater than 1; the non-intersecting intervals avoid that problem in this case. The non-intersection also implies that ρ must have exactly four transitions, and of course ρ only takes on the values 0 and 1.

Geometrically, γ_1 and γ_2 are two XYZ vectors in $\mathbb{R}^3$, on the boundary of Ω_I. They can be added to produce a third vector

$$\gamma = \gamma_1 + \gamma_2. \tag{13}$$

By linearity, the reflectance spectrum ρ must be in the metamer set of γ. Since ρ has more than two transitions, the Optimal Colour Theorem implies that γ cannot be on the boundary of Ω_I. Since ρ is a physically possible object colour, however, γ must belong to Ω_I, and is therefore a non-optimal, interior colour, for which Equation (13) provides a four-transition 0-1 function.

8
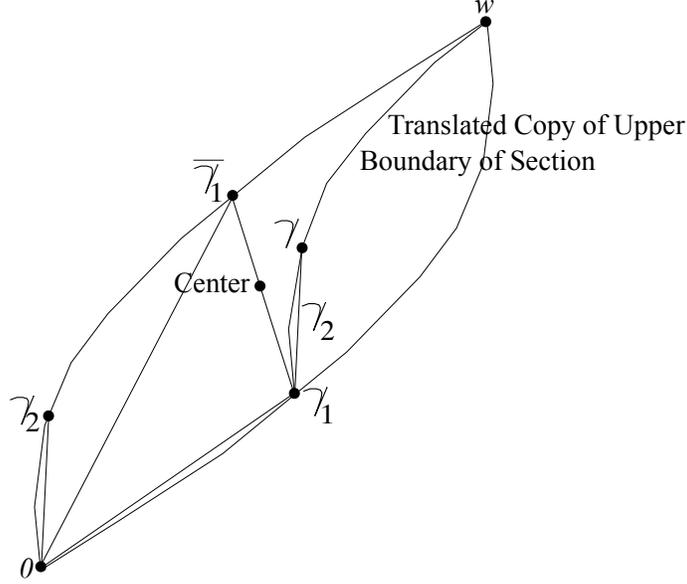
Figure 5: Disjoint Optimal Spectra that Sum to γ

3.2   Finding Disjoint Optimal Colours

Rather than starting with optimal colours, this section starts with a non-optimal colour γ in the interior of Ω_I, and shows the existence of a four-transition 0-1 metamer for γ by finding disjoint optimal colours γ_1 and γ_2 that sum to γ. Begin by constructing the plane $\mathcal{P}$ which contains γ, the origin, and the white point. (If γ is on the neutral line joining the origin and the white point, then there are many such planes; choose any one.) Figure 5 shows the section $\Omega_I \cap \mathcal{P}$, with the point γ in its interior.

The desired vectors γ_1 and γ_2 must satisfy Equation (13). They are optimal, so they are on the boundary of Ω_I, and will also appear on the boundary of the section. They will be chosen to be disjoint; one will be on the lower boundary of the section, and the other will be on the upper boundary. A geometric test will tell us whether a given γ_1 on the lower boundary is disjoint from a γ_2 on the upper boundary. Start with γ_1 as shown in Figure 5, and construct the complementary colour $\bar{\gamma}_1$. Considering the colours as vectors, we must have

$$w = \gamma_1 + \bar{\gamma}_1, \tag{14}$$

while considering the colours as reflectance spectra gives

$$\rho_w = \rho_1 + \bar{\rho}_1, \tag{15}$$

where ρ_w is the ideal white spectrum, which is identically 1.

Now consider any point γ_2 that is on the upper boundary between 0 and $\bar{\gamma}_1$. Then γ_2 is also optimal, and, as shown earlier, its support must be a proper subset of the support of $\bar{\gamma}_1$. The sum $\gamma_1 + \gamma_2$ is therefore a four-transition, 0-1 function which is contained in the section. As γ_2 varies from 0 to $\bar{\gamma}_1$, the set of sums $\gamma_1 + \gamma_2$ will trace out a translated copy of the part of the upper boundary between 0 and $\bar{\gamma}_1$; the copy starts at γ_1 and ends at w, given by $\gamma_1 + \bar{\gamma}_1$. Any point on this copy can be written as a four-transition, 0-1 function.

The final step in the construction is now straightforward: if one pictures the upper boundary sliding along the lower boundary, it is clear that the entire section will be swept out, and in particular any desired γ inside the section will sooner or later be reached. When a desired γ is reached, then choose as γ_1 the point where the translated copy originates on the lower boundary, and let γ_2 be the difference of γ and γ_1. These optimal colours are disjoint and sum to γ as was to be shown.

While proving the existence of a four-transition, 0-1 metamer, the construction should also make it clear that these metamers are not unique. A simple example of non-uniqueness would involve the center point itself, or indeed any colour on the line segment between 0 and w. There are many choices of a section for these cases, and any section will give a pair of optimal colours that do not occur in the other sections. The sums of these pairs will therefore also be distinct as reflectance spectra, although they all produce the same XYZ. For a point not on the central segment, one could choose a section that contained 0 and that point, but not w. Since many of the sums of pairs of optimal colours on the boundary of that section would still be within the colour solid, they would provide alternate four-transition 0-1 metamers for many non-central colours. In general, then, four-transition 0-1 reflectance spectra would not be unique for non-optimal colours.

4   Discussion

The four-transition 0-1 functions described here stemmed from a paper of Logvinenko, Funt, and Godau,[11] who used five-transition 0-1 functions (which would have six transitions by our way of counting) for reflectance spectra, in an algorithm to construct metamer mismatch bodies[2] (MMBs) for imaging sensors. MMBs involve two illuminants, and a six-dimensional object that might be called a double object-colour solid, originally introduced by Eugene Allen,[12] was constructed, to account for both illuminants simultaneously. The boundary of a six-dimensional body is five-dimensional, motivating their choice of five transitions. While a sound modeling decision, questions naturally arose about whether, as in Schrödinger's original result, certain forms were geometrically inevitable, or could be shown mathematically to be adequate. The results of this paper are a first step in addressing these questions; it is hoped that further development along these lines will be useful in applications.

5   Summary

In analogy to the two-transition 0-1 Schrödinger form for optimal colours, the current paper has shown that every non-optimal colour is metameric to the colour produced by a four-transition 0-1 reflectance spectrum. Although the Schrödinger form is unique for optimal colours, the four-transition form is not unique for non-optimal colours. The demonstration of these results used the convexity and symmetry of object-colour solids, and the concomitant convexity and symmetry of central sections of those solids, to express an arbitrary non-optimal colour as the sum, in CIE XYZ space, of two optimal colours with non-intersecting support. The corresponding algebraic sum of the optimal reflectance spectra is then a four-transition 0-1 spectrum as desired.

References

1. Deane B. Judd, "The 1931 I. C. I. Standard Observer and Coordinate System for Colorimetry," JOSA, Vol. 23, pp. 359-374, October 1933.

2. G. Wyszecki and W. S. Stiles, *Color Science: Concepts and Methods, Quantitative Data and Formulae, 2ⁿᵈ ed.*, John Wiley & Sons, 1982.

3. Erwin Schrödinger. "Theory of pigments of greatest lightness," translated by Rolf G. Kuehni with a technical introduction by Michael H. Brill, accessed Feb 2018 from http://www.iscc-archive.org/resources/translations.php (original article: "Theorie der Pigmente von grösster Leuchtkraft," Annalen der Physik 4, 62, 1920, pp. 603-622).

4. David L. MacAdam. "The Theory of the Maximum Visual Efficiency of Colored Materials," JOSA, Vol. 25, 1935, pp. 249-252.

5. Paul Centore, "Non-Metamerism of Boundary Colors in Multi-Primary Displays," Journal of the Society for Information Display, Vol. 20, Issue 4, pp. 214-220, Apr 2012.

6. Paul Centore, *The Geometry of Colour*, 221 pages, 2017.

7. A. D. Logvinenko, "An object-color space" (see footnote 1), Journal of Vision (11):5, pp. 1-23, 2009.

8. Alexander D. Logvinenko & V. I. Levin, *Foundations of Colour Science*, unpublished manuscript.

9. Paul Centore, "A Zonohedral Approach to Optimal Colours," Color Research & Application, Vol. 38, No. 2, pp. 110-119, April 2013.

10. David L. MacAdam. *Color Measurement: Theme and Variations* (see Sect. 7.5), Springer Series in Optical Science, Vol. 27, Springer-Verlag, 1981.

11. Alexander D. Logvinenko, Brian Funt, & Christoph Godau, "Metamer Mismatching," IEEE Transactions on Image Processing, Vol. 23, No. 1, pp. 34-43, January 2014.

12. Eugene Allen, "Some New Advances in the Study of Metamerism: Theoretical Limits of Metamerism; An Index of Metamerism for Observer Differences," Color Engineering, Vol. 7, No. 1, pp. 35-40, January-February 1969.