Non-Metamerism of Boundary Colours in Multi-Primary Displays

Paul Centore*

Abstract

A control sequence gives the intensities of the primaries for a pixel of a display device. The display gamut, i.e. the set of all the colours that a display can produce, is a zonohedral subset of CIE XYZ space, and contains both boundary and interior colours. Displays with four primaries or more exhibit metamerism, in which different control sequences produce colours that appear identical to an observer. This paper shows mathematically that, provided no three primaries are linearly dependent, metamerism can only occur for interior colours. When there are four or more primaries, metamers can always be found for interior colours. A colour on the gamut boundary, by contrast, is only produced by a unique control sequence. The proof used for displays can be extended to object-colour solids, to show that optimal colours, which are on the boundary of an object-colour solid, have unique reflectance functions.

Keywords — multi-primary, metamerism, zonohedron, control sequence, gamut

1 Introduction

Display devices such as computer monitors and televisions typically produce three primaries: red, green, and blue. A primary is a coloured light source whose intensity varies, but whose relative spectral power density does not. Combining the primaries at differing intensities produces a wide gamut of colours. Recently, multi-primary displays, involving more than three primaries, have been introduced, to produce an even wider gamut. The list of coefficients in a combination of primaries is called a

*Paul Centore lives at 62 Chicago Ave., Groton, CT 06340; e-mail: centore@99main.com

control sequence. Each coefficient in a control sequence is between 0 (i.e., a primary is not activated at all) and 1 (i.e., a primary is activated at maximum intensity).

In 1931, the Commission Internationale de l'Éclairage (CIE) defined a standard observer,[1] who converts a colour stimulus to three perceptual coordinates: X, Y, and Z. Two colour stimuli, such as two combinations of primaries, appear identical to the standard observer if and only if their XYZ coordinates are equal. The display gamut is the subset of XYZ space, consisting of the XYZ coordinates for all possible control sequences. When written in XYZ coordinates, each primary in a display becomes a vector. The gamut is the Minkowski sum of the primaries, in three-dimensional XYZ space, and is thus a zonohedron. Every zonohedron is a closed, convex set. As long as there are at least three independent primaries, the zonohedron has a non-empty interior. Furthermore, any point of a zonohedron that is not an interior point is a boundary point.

Human colour vision is three-dimensional, so the control sequence for any colour in the gamut is unique when there are only three linearly independent primaries. Adding more primaries expands the device's colour gamut, but introduces *metamerism*: a colour in the gamut might result from many different combinations of the primaries, instead of from a unique combination. Two control sequences are metameric if and only if their XYZ vectors are equal.

This paper will prove mathematically that, as long as no three primaries are linearly dependent, metamerism cannot occur on the gamut's boundary. In other words, a colour on the boundary of the display gamut can be produced by one, and only one, control sequence. An interior point, by contrast, is produced by a unique control sequence only when there are three primaries. If there are four or more primaries, then any interior point can be produced by multiple control sequences.

The proof involves interpreting the zonohedron's support planes as the maximum of a linear functional. The proof applies not only to display gamuts, but also to object-colour solids, which are zonohedral combinations of monochromatic stimuli. The boundary of an object-colour solid consists of so-called optimal colours, and the proof can be extended to show that an optimal colour results from a unique reflectance function.

A statement of the uniqueness result occurred in a 2004 paper by Kanazawa *et al*,[2] who used it in their spherical average construction. No proof was given, however. The proof in the current paper fills this gap, and provides a rigorous basis for their construction.

2 Uniqueness and Non-Uniqueness of Control Sequences

2.1 Geometry of Display Gamuts

A display device combines a limited set of primaries to produce a wide gamut of colours. A primary is a light source of fixed relative spectral density, whose intensity can vary from 0 to some maximum, which can be denoted 1. Geometrically, primaries and their combinations can be represented by CIE XYZ coordinates, which can be viewed as vectors in the positive octant of $\mathbf{R}^3$. The set of all possible combinations of primaries forms a solid, called the display gamut, in the positive octant. Suppose that N primaries are available. Denote the XYZ vector of the i^{th} primary by $\mathbf{v_i}$. The display gamut, G, consists of all possible combinations of primaries, at all intensities from 0 to 1:

$$G = \left\{ \sum_{i=1}^{N} \alpha_i \mathbf{v_i} \middle| \alpha_i \in [0,1] \, \forall i \right\}. \tag{1}$$

The XYZ coordinate vector of a combination of primaries, such as appear in Equation (1), is the sum, in the vector space sense, of the XYZ coordinates of the individual primaries, multiplied by the appropriate α_i. A sequence of N coefficients $\alpha_i, i = 1 \ldots N$, all of which are between 0 and 1, is called a *control sequence.*

Equation (1) shows that the gamut G is the zonohedron generated by the N primary vectors in the positive octant. A zonohedron is the Minkowski sum of a set of vectors. The Minkowski sum, or vector sum, of two sets, $\mathbf{A}$ and $\mathbf{B}$, in $\mathbf{R}^n$, is defined as

$$\mathbf{A} \oplus \mathbf{B} = \{a + b | a \in \mathbf{A}, b \in \mathbf{B}\}. \tag{2}$$

More concretely, the Minkowski sum of $\mathbf{A}$ and $\mathbf{B}$ is the volume (or area) that $\mathbf{B}$ sweeps out when its tail can be located at any point in $\mathbf{A}$. The Minkowski sum is both commutative and associative.

The Minkowski sum of two vectors is the parallelogram swept out by letting either vector slide along the other. The Minkowski sum of a parallelogram and another vector, not in the same plane, is the parallelepiped swept out by sliding the parallelogram along that vector. Figure 1 shows the Minkoswki sum of a parallelepiped and another vector in $\mathbf{R}^3$. The parallelepiped generated by $\{\mathbf{v_1}, \mathbf{v_2}, \mathbf{v_3}\}$ appears on the left. On the right is the zonohedron of $\{\mathbf{v_1}, \mathbf{v_2}, \mathbf{v_3}, \mathbf{v_4}\}$. To make the figure on the right, place a copy of $\mathbf{v_4}$ at each vertex of the figure on the left, and then take the convex hull; alternately, slide the parallelepiped on the left along the vector $\mathbf{v_4}$.

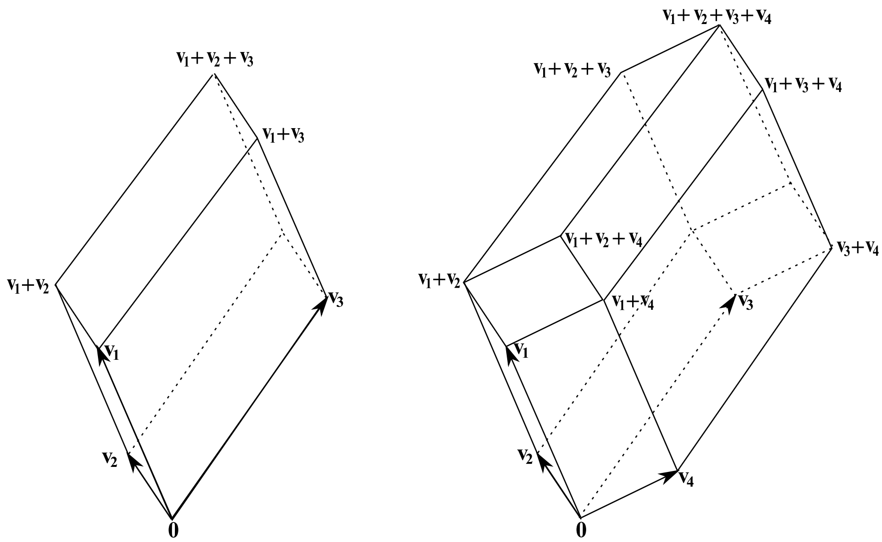

Figure 1: Construction of Zonohedron

Zonohedra are convex and two-fold rotationally symmetric. A zonohedron's faces are all parallelograms, provided that no three generating vectors are coplanar, and no generating vector is a multiple of another generating vector. These two statements are subsumed in the equivalent single statement that no three generating vectors are linearly dependent.

Expression (1) shows that there can be some redundancy, or *metamerism*, in the display gamut. Metamerism in the context of display devices occurs when two different control sequences produce the same XYZ tristimulus values. From a practical point of view, metamerism should usually be avoided in displays, because individuals who differ slightly from the CIE standard observer might see two metameric colours as different, when they should appear identical. Metamerism cannot occur when there are only three (linearly independent) primaries. XYZ space can be seen as a vector space, in which Equation (1) gives linear combinations of the primaries. When there are only three primaries, they form a basis for the vector space, so any point can be written as only one linear combination of the primaries.

2.2 Proofs of Uniqueness and Non-Uniqueness

As a convex set, a zonohedron can be partitioned into interior points and boundary points. The main result of this paper is that metamerism cannot occur for boundary points, although it can always occur for interior points, when there are four primaries or more. Stated otherwise: when a point is on the boundary of the zonohedral gamut, it is generated by a unique control sequence. Kanazawa[2] states this fact, but does not prove it. We will adapt the methods used by McMullen[3] to give a proof.

Theorem 1. *If no three primaries are linearly dependent, then any point P on the boundary of a display gamut G, generated by an arbitrary number of primaries, is produced by a unique control sequence.*

Proof. The proof uses the characterization of a convex set (such as the display gamut) as the intersection of all the half-spaces containing it.[4] A support plane of the gamut G is a plane that contains at least one point of G, but no interior point of G. Support planes intersect the boundary of G; the intersection with the boundary can be a vertex, an edge, or a face. In this proof, only support planes that intersect faces will be needed. G (or at least, its interior) is in a half-space defined by any support plane. Furthermore, a support plane S is as close as possible to, but does not actually intersect, the interior of G, when compared to parallel planes. Since the zonohedral gamut starts at the origin, and reaches the support plane, but does not reach any further, therefore any combinations of generating vectors that reach the support plane must be maximal in a sense that we will make precise.

In a general vector space V, such as CIE XYZ space, there is no inner product with which to measure the distance of a hyperplane (such as a support plane), from the origin. However, as shown in Chapter 2 of a textbook by Lay,[5] every hyperplane is a level set of some linear functional, and conversely, all the level sets of a (non-zero) linear functional are parallel hyperplanes. Therefore, given a support plane S, that does not contain the origin, we can find a linear functional f_S on V, such that $f_S(S) = 1$. For any k, kS is a plane parallel to S, and $f_S(kS) = k$. We can think of f_S as defining a distance from the origin to any plane parallel to S. If the support plane contains the origin, then let f_S be any linear functional that is 0 on the support plane, and positive everywhere on the zonohedron, which is always on one side of the support plane.

Now choose a point P that is on a face of the gamut boundary, but not on an edge or vertex. We will first show that there is a unique control sequence for P, when P is interior to a face, and then proceed to the case where P is on a vertex or edge. Let S be the support plane through P. Since P is interior to a face F, it follows that S is unique, and contains F. Define f_S as was explained in the previous paragraph. Translate S to a parallel plane S_0, that intersects the origin. S_0 divides the generating vectors at the origin into three groups: those for which f_S is negative, those for which it is positive, and those for which it is zero. Call the first group the negative generators, the second group the positive generators, and the third group

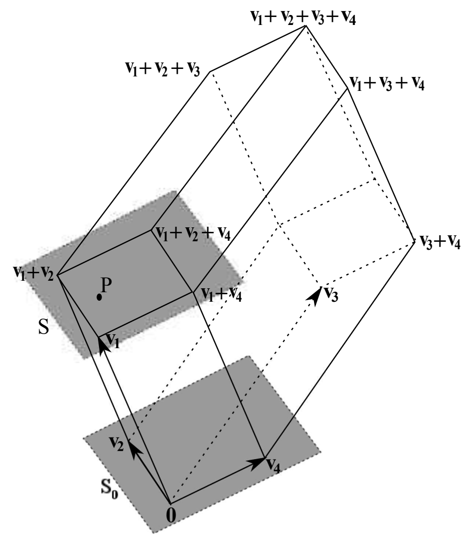

Figure 2: Translated Support Plane, Separating Generating Vectors

the zero-level generators.

Figure 2 shows an example. The face F containing P is the parallelogram with vertices $\mathbf{v_1}$, $\mathbf{v_1} + \mathbf{v_2}$, $\mathbf{v_1} + \mathbf{v_4}$, and $\mathbf{v_1} + \mathbf{v_2} + \mathbf{v_4}$. The support plane S contains translated copies of $\mathbf{v_2}$ and $\mathbf{v_4}$. S_0, given by translating S to the origin, contains the original $\mathbf{v_2}$ and $\mathbf{v_4}$. $\mathbf{v_2}$ and $\mathbf{v_4}$ are therefore the zero-level generators. The functional f_S is positive for vectors pointing away from S_0 and towards S, so $\mathbf{v_1}$ is a positive generator. Since $\mathbf{v_3}$ points towards the other side of S_0, it is a negative generator.

The negative and positive groups lie on different sides of the plane through the origin. We will show that the coefficients in the control sequence must be 1 for the positive generators, and 0 for the negative generators. Since the zonohedron is on one side of the support plane, but as close as possible to it, therefore the support plane represents the maximum non-negative value that f_S attains on the zonohedron. Denote a point $\mathbf{v}$ of the zonohedron by

$$\mathbf{v} = \sum_{i=1}^{N} \alpha_i \mathbf{v_i}. \tag{3}$$

By linearity,

$$f_S(\mathbf{v}) = \sum_{i=1}^{N} \alpha_i f_S(\mathbf{v_i}). \tag{4}$$

The positive generators contribute positive values of f_S to Expression (4), while the negative generators contribute negative values. Since all quantities are fixed in Expression (4) except α_i, the expression is maximized by making α_i as large as possible, which means setting it to 1, for positive generators. Similarly, the contributions of the negative generators must be minimized, so their coefficients are set as small as possible, which means 0. This maximum gives unique coefficients for the positive and negative generators. Denote the sum of the positive generators, all with coefficients 1, by $\mathbf{w}$. By construction, $\mathbf{w}$ is a point in the original S, and the vector $P - \mathbf{w}$ can be translated to lie in S_0.

The remaining generators, if any, must be zero-level generators in S_0. Since it was assumed that no three primaries are linearly dependent, there can be at most two zero-level generators, and the set of zero-level generators is itself linearly independent. P is the sum of $\mathbf{w}$ and a positive linear combination of the (at most two) zero-level generators. Since there are no more than two zero-level generators, and since S_0 is a two-dimensional vector subspace of V, therefore $P - \mathbf{w}$ has unique coefficients as a linear combination of the zero-level generators. These same unique coefficients apply when P is written as

$$P = \mathbf{w} + (P - \mathbf{w}), \tag{5}$$

because the bases for the positive generators and zero-level generators are distinct.

In summary, then, any P, that is interior to a face, has unique coefficients for all positive, negative, and zero-level generators, which partitions the set of generating vectors into distinct subsets. The coefficients give a unique control sequence for P. The positive generators can be seen as pushing a hyperplane away from the origin, through the zonohedral gamut, as far as possible. The maximum distance is achieved when the hyperplane becomes a support plane. The requirement to reach a maximum guarantees uniqueness. Once the plane has become a support plane, a linear combination of the zero-level generators defines the position of P in the plane, and the uniqueness of this linear combination follows from linear algebra. Furthermore, a vertex can be written as the sum of all generating vectors which are on one side of a plane through the origin. In fact, the converse is also true (though not proven here). Draw a plane through the origin, and add up all the generating vectors on one side of it: then this sum is a vertex.

In the example in Figure 2, the only positive generator is $\mathbf{v_1}$, so we have simply $\mathbf{w} = \mathbf{v_1}$. Furthermore, there are exactly two zero-level generators, $\mathbf{v_2}$ and $\mathbf{v_4}$. The unique control sequence of P is

$$P = \mathbf{v_1} + \frac{1}{2}\mathbf{v_2} + \frac{1}{4}\mathbf{v_4}. \tag{6}$$

The foregoing constructions allow a finer description of any parallelogram face F. The vector $\mathbf{w}$ gives a distinguished vertex of F, which we call the originating vertex. The edges of F that originate at $\mathbf{w}$ will be called originating edges. The originating edges are translates of two generating vectors, $\mathbf{v_a}$ and $\mathbf{v_b}$. The other two edges are terminating edges, and the vertex opposite $\mathbf{w}$ is the terminating vertex. A parallelogram's vertices can then always be written in a unique way as $\mathbf{w}$, $\mathbf{w}+\mathbf{v_a}$, $\mathbf{w}+\mathbf{v_b}$, and $\mathbf{w}+\mathbf{v_a}+\mathbf{v_b}$.

We will now show that any gamut vertex V has a unique control sequence. By way of contradiction, assume that V has two control sequences:

$$V=\sum_{i=1}^{N}\alpha_i\mathbf{v_i},\ \text{where } \alpha_i\in[0,1], \tag{7}$$

and

$$V=\sum_{i=1}^{N}\beta_i\mathbf{v_i},\ \text{where } \beta_i\in[0,1]. \tag{8}$$

Construct V_{min}, given by choosing the smaller of the two coefficients for each generating vector:

$$V_{\text{min}}=\sum_{i=1}^{N}\min(\alpha_i,\beta_i)\mathbf{v_i}. \tag{9}$$

The difference $V-V_{\text{min}}$ has two expressions, one involving α's and the other involving β's. Denote these two expressions by V_α and V_β :

$$V-V_{\text{min}} = \sum_{i=1}^{N}(\alpha_i-\min(\alpha_i,\beta_i))\mathbf{v_i}=V_\alpha \tag{10}$$

$$= \sum_{i=1}^{N}(\beta_i-\min(\alpha_i,\beta_i))\mathbf{v_i}=V_\beta. \tag{11}$$

Now consider the following three vectors:

$$V_{\text{min}}, \tag{12}$$

$$V_{\text{min}}+V_\alpha, \tag{13}$$

$$V_{\text{min}}+V_\alpha+V_\beta. \tag{14}$$

The first one, V_{min}, is in the gamut, because all the coefficients are between 0 and 1, as can be seen from Equation (9). The second one, $V_{\text{min}} + V_\alpha$, is just the original vertex V, as can be seen from Equation (10), so it is also in the gamut. Finally, Equations (10) and (11) can be used to write the third vector as:

$$V_{\text{min}} + V_\alpha + V_\beta = \sum_{i=1}^{N} (\alpha_i + \beta_i - 2\min(\alpha_i, \beta_i)).\mathbf{v_i} \tag{15}$$

Since α_i and β_i are both between 0 and 1, the coefficients in Equation (15) are between 0 and 1 as well, so the third vector is also in the gamut. Therefore each of the three vectors is in the gamut. Furthermore, Equations (10) and (11) show that $V_\alpha = V_\beta$, so the third vector can be rewritten $V_{\text{min}} + 2V_\alpha$. Written in this way, it becomes clear that the three vectors in Equations (12) through (14), all of which are in the gamut, lie on a straight line, with the vertex V in the middle. Since the gamut is convex, this is a contradiction. The only resolution is to conclude that V_α is the zero vector, which is to say, that the representation of V in Equation (7) is unique, as desired.

Finally, we will show that all points on the edges of the zonohedral gamut have unique control sequences. Let P be a point on an edge, E, that joins two vertices, V_1 and V_2. In a convex solid such as the zonohedral gamut, an edge can only occur as a line segment joining one pair of vertices, so V_1 and V_2 are unique. From the construction in Figure 1, V_1 and V_2 differ by a generating vector, $\mathbf{v_g}$. Switching the numbering of V_1 and V_2 if necessary, we obtain

$$V_2 = V_1 + \mathbf{v_g}. \tag{16}$$

Since P is on the line joining V_1 and V_2, it is possible to write

$$P = V_1 + k\mathbf{v_g}, \tag{17}$$

where $k \in [0, 1]$. Now let P be given by some control sequence:

$$P = \sum_{i=1}^{N} \alpha_i \mathbf{v_i}, \text{ where } \alpha_i \in \{0, 1\}. \tag{18}$$

We will show that, the coefficient α_g of $\mathbf{v_g}$ in *any* control sequence for P, is in fact equal to k. Substitute Equation (18) into Equation (17):

$$\sum_{i=1}^{N} \alpha_i \mathbf{v_i} = V_1 + k\mathbf{v_g} \tag{19}$$

$$V_1 = \sum_{i \neq g} \alpha_i \mathbf{v_i} + (\alpha_g - k)\mathbf{v_g}. \tag{20}$$

Since V_1 is a vertex, its control sequence is unique. In addition, it follows from Equation (16) that the coefficient of $\mathbf{v_g}$ is 0. Therefore, in any control sequence for P, it must be that $\alpha_g = k$, so α_g is unique. But more can be inferred from Equations (17) and (20): if there were two distinct control sequences for P, then there would be two distinct control sequences for the vertex V_1. Since it has already been shown that vertices have unique control sequences, it follow that control sequences are also unique for a point P on an edge of the gamut.

In summary, we have shown that vertices, points on edges, and points in the interior of the gamut's faces, all have unique control sequences. In other words, any colour on the boundary of the display gamut has a unique control sequence, as was to be shown. ■

When there are only three independent primaries, the zonohedral gamut is a parallelepiped. Since the three primaries are independent, they form a basis for XYZ space, so any point, in particular any point inside the gamut, has a unique control sequence. In the typical three-primary display, then, control sequences are unique for both the interior and the boundary of the gamut.

When there are more than three primaries, boundary points have unique control sequences, but interior points do not:

Theorem 2. *Suppose a display device has four or more primaries, which generate a display gamut, G, with non-empty interior. Then a point P in the interior of G can always be produced by multiple control sequences.*

Proof. We will first show that P can be written with a control sequence whose coefficients are all greater than 0 and less than 1. Draw the line L from the origin (where every control sequence coefficient is 0) to P. Since G is convex, L is contained in G. Since P is in the interior of G, there exists a sphere σ_P around P that is completely contained within G. Extend L to a longer line L', that continues past P, to a new point P', that is inside σ_P. There is a positive ϵ such that $L' = (1+\epsilon)L$. Since P' is within G, there is some set of coefficients, α'_i, such that

$$P' = \sum_{i=1}^{N} \alpha'_i \mathbf{v_i}. \tag{21}$$

Every primed coefficient is between 0 and 1, and may possibly take on the values 0 or 1. As vectors, $P' = (1+\epsilon)P$, so

$$P = \sum_{i=1}^{N} \frac{\alpha'_i}{1+\epsilon} \mathbf{v_i}. \tag{22}$$

Because of the denominator, all the coefficients for P in Equation (22) are now strictly less than 1.

A set of coefficients for P, that are not only all strictly less than 1, but also all strictly greater than 0, is found similarly. Draw the line M that connects $F = \sum \mathbf{v_i}$ to P. The point F is the point on G that is farthest from the origin. Similarly to the previous step, extend M to a longer line M'', that continues past P, to a new point P'', that is in σ_P. There is some positive δ such that

$$P'' - F \;=\; (1+\delta)(P - F). \tag{23}$$

Now apply Equation (22) to P'', instead of P, to conclude that there exist coefficients α_i'', such that

$$P'' = \sum_{i=1}^{N} \alpha_i'' \mathbf{v_i}, \tag{24}$$

where every α_i'' is strictly less than 1, and some α_i'''s might be 0. Substitute the vector space expressions for F and P'' into Equation (23), and rearrange to get

$$P = \sum_{i=1}^{N} \frac{\alpha_i'' + \delta}{1+\delta} \mathbf{v_i}. \tag{25}$$

Since every α_i'' is strictly less than 1, it follows that every coefficient in Equation (25) is also less than 1. Since every α_i'' is greater than or equal to 0, the numerator insures that every coefficient in Equation (25) is strictly greater than 0. Equation (25) therefore presents a control sequence for an interior point P, such that every coefficient is strictly between 0 and 1, as desired.

To finish proving the theorem, assume, by way of contradiction, that there is a unique control sequence, α_{iP}, such that

$$P = \sum_{i=1}^{N} \alpha_{iP} \mathbf{v_i}. \tag{26}$$

P is an interior point of G, so the previous argument, combined with the assumed uniqueness, implies that $0 < \alpha_{iP} < 1$, for all i. Since there are more than three primaries, in a three-dimensional space, there exists a set of coefficients β_i, not all 0, such that

$$\sum_{i=1}^{N} \beta_i \mathbf{v_i} = \mathbf{0}. \tag{27}$$

Equation (27) will hold even if all the β's are multiplied by an arbitrarily small constant γ. Multiplying Equation (27) by γ, and adding Equation (26) gives

$$P = \sum_{i=1}^{N} (\alpha_{iP} + \gamma\beta_i)\mathbf{v_i}. \tag{28}$$

The original control coefficients, α_{iP}, are all strictly between 0 and 1. By choosing γ small enough, the coefficients $\alpha_{iP} + \gamma\beta_i$ are also strictly between 0 and 1, and therefore Equation (28) defines a point in G. The coefficients in Equations (26) and (28) are different, however, proving that the control sequence in Equation (26) cannot be unique. ■

Theorem 2 shows that metamerism is inevitable with multi-primary displays. As soon as a fourth primary is added, points inside the gamut have multiple representations. Metamerism can therefore not be avoided by choosing primaries judiciously. Neither is it possible that metamerism will be a problem only with some colours inside the gamut: Theorem 2 shows that *every* interior colour can be produced by multiple control sequences. By contrast, colours on the boundary gamut are better behaved. It is in fact impossible to use different representations for a boundary colour, because Theorem 1 shows that control sequences are unique on the gamut boundary.

3 Discussion

As mentioned, the uniqueness result was stated without proof by Kanazawa *et al.* They also omitted the technical requirement that no three primaries are linearly independent. They used the result in their spherical average construction. This construction expresses a point in the gamut's interior as a weighted sum of points on the gamut's boundary. Since the boundary control sequences are unique, the same weighted sum, applied to the unique control sequences of the boundary points, gives a unique control sequence for points inside the zonohedron. Furthermore, the averaging process guarantees that the assignment of control sequences is continuous. Their construction avoids metamerism and spurious discontinuities in multi-primary displays. By filling a technical gap, the current proof provides a rigorous basis for Kanazawa's work.

Another display application is the matrix-switching method of Ajito *et al.*[6] The matrix-switching method partitions a display gamut into irregular pyramids, each of whose bases is a boundary parallelogram, and the common apex of which is the origin. The assignment method involves a different matrix for each pyramid. Their construction can be more directly seen, though perhaps not easily calculated, by invoking the uniqueness of boundary control sequences. Given a point P inside the gamut, draw the ray from the origin through P. That ray crosses the gamut boundary at some point C. By convexity, this crossing point is unique. Because P and C are on the same line through the origin, there is a positive constant k, less than 1, such that $P = kC$. The control sequence for the boundary point C is unique. Defining the control sequence for P to be k times the control sequence for C specifies a single sequence for P, thus avoiding on-screen metamerism, which was one of the goals. In fact, this assignment method gives the same result as Ajito's matrix-switching method, but with a simpler interpretation.

The current proof is believed to be the first published proof of control sequence uniqueness on a gamut boundary, as well as the first to state the technical requirement that no three primaries are linearly independent. A proof of a similar statement, about the object-colour solid, has been claimed by A. Logvinenko. In a footnote to a 2009 paper,[7] he wrote "Although this intuitively clear statement [that there is no metamerism on the boundary of the object-colour solid] has been made by several authors before, ... the formal proof has not been known until recently." A reference is then given to an unpublished manuscript, *Foundations of colour science*. It is likely that techniques of analyzing the object-colour solid would shed light on display gamuts as well, so it would be interesting to compare the approaches once they are both published.

For many applications, it would be desirable for the display gamut and the object-colour solid to be identical subsets of XYZ space. Since displays have a finite number of primaries, however, while the object-colour solid is generated by the infinite set of monochromatic vectors in the spectrum locus, the two subsets will never be identical. Nevertheless, their structures are similar. A recent paper of Centore[8] approximates the object-colour solid as a zonohedron. In that paper, the generating vectors were XYZ vectors for monochromatic spectral densities. In practice, a finite set of such densities is used, while in theory the set is infinite. As the finite set becomes infinite in the limit, the zonohedra corresponding to the finite sets approach a set called a zonoid. This zonoid has a smooth surface, with no edges or faces, although there are vertices at ideal black and ideal white. A display gamut, by contrast, is genuinely a zonohedron, because the set of primaries is always finite. Thus a display gamut has a finite set of parallelogram faces.

A colour on the boundary of the object-colour solid is called an *optimal colour*. The Optimal Colour Theorem states that the reflectance function (over the visible spectrum) of an optimal colour takes on only the values 0% and 100%, with at most two transitions between those values. By applying the methods of Theorem 1 to object-colour solids, and using the empirical finding that no three vectors in the spectrum locus are linearly dependent, it follows that there is a *unique* reflectance function for each optimal colour. Centore discretized reflectance functions over the visible spectrum, and constructed a zonohedral approximation to the object-colour solid. Optimal colours were taken to be the vertices of that zonohedron. The Optimal Colour Theorem then followed from the convex, cyclic form of the spectrum locus, and the fact that a zonohedron's vertices are sums of primaries at maximum activation. Rather than just taking vertices as optimal colours, any boundary point of the zonohedron could have been taken as an optimal colour; this interpretation was not needed because the distinction between vertices and other boundary points disappears in the zonoidal limit. In the case of display gamuts, non-vertex boundary points must be considered—in fact, they make up far more of the boundary than the few vertices. Thus the detailed proof of Theorem 1 was necessary.

Stiles and Wyszecki, along with other researchers, approached the problem of metamerism and uniqueness from a computational point of view.[9] Using a Monte Carlo simulation, as well as linear programming and the central limit theorem, Stiles and Wyszecki calculated how many metamers there were for a particular colour in the object-colour solid. Since the object-colour solid and a display gamut are similarly shaped subsets of XYZ space, their methods would work equally as well for estimating how many control sequences could produce a particular colour in a display gamut. An overall finding was that there were many metamers for colours near the "center" of the solid, and progressively fewer farther away from the center, terminating with a unique metamer on the boundary.

Though intuitively correct, their results should be interpreted with caution. The reason is that, when counting "how many" metamers, they assigned uniform weightings to the wavelengths in the visible spectrum, and to the possible values for the reflectance factor for each wavelength. While this assignment seems natural enough, it is not clear that there is any physical or perceptual justification for it. For a multi-primary display system, however, it might be possible to find a weighting with a practical basis. For example, one could count the number of subsets of primaries that could be used to generate a given gamut colour. In a four-primary display, some colours require that the coefficient of $\mathbf{v_4}$ be non-zero, so those colours would have fewer metamers than those which could be generated from the first three primaries.

4 Conclusion

This paper fills a technical gap, by presenting a mathematical proof of the statement that metamerism does not occur on a display gamut boundary. Similar statements have been made that optimal colours, on the boundary of the object-colour solid, have unique reflectance functions. Although empirical evidence for these statements has been found, it is believed that no proof has been published before now. In addition, this paper is believed to be the first to state the technical requirement that no three primaries are linearly dependent in CIE XYZ space. A related result, also proven here, is that metamerism is inevitable on a gamut's interior when there are more than three primaries. All the proofs use a geometric approach, involving the zonohedral structure of the display gamut.

1 D. B. Judd, "The 1931 I. C. I. Standard Observer and Coordinate System for Colorimetry," *J Opt Soc Am* **23**, 359-374 (1933).
2 H. Kanazawa, M. Mitsui, M. Yamaguchi, H. Haneishi, and N. Ohyama, "Color Conversion for Multi-primary Displays Using a Spherical Average Method," *Proc CIC*, 65-69 (2004).
3 P. McMullen, "On Zonotopes," *Trans Am Math Soc* **159**, September, 91-109 (1971), Sect. 2.
4 R. V. Benson, *Euclidean Geometry and Convexity* (New York: McGraw-Hill Book Company, 1966), Sect. 2.8.
5 S. R. Lay, *Convex Sets and Their Applications* (New York: Dover Publications, Inc., 2007).
6 T. Ajito, K. Ohsawa, T. Obi, M. Yamaguchi, and N. Ohyama, "Color Conversion Method for Multiprimary Display Using Matrix Switching," *Opt Rev* **8**, No. 3, 191-197 (2001).
7 A. D. Logvinenko, "An object-color space," *Journal of Vision* **9(11):5**, 1-23, 2009.
8 P. Centore, "A Zonohedral Approach to Optimal Colours," accepted for publication in *Color Research and Application*, 2012.
9 G. Wyszecki and W. S. Stiles, *Color Science: Concepts and Methods, Quantitative Data and Formulae, 2nd ed.* (New York: John Wiley & Sons, 1982), Sect. 3.8.4.

