Avoiding on-screen metamerism in *N*-primary displays

Michael H. Brill
James Larimer

Abstract — The present paper describes a method for using more than three primaries in an additive-primary display. The method ensures that each tristimulus specification can be produced in no more than one way, even if a non-singular filter (*i.e.,* one that does not reduce the dimensionality of color-matching space) is interposed between the screen and the viewer. Starting with *N* primaries, the method uses only three at a time, but these may be composites – fixed linear combinations of the original *N*. As further insurance against on-screen metamerism, a criterion on the primary spectra, based on the Binet-Cauchy theorem, ensures that a triad of primaries keeps its right/left-handed chromaticity ordering when a filter is interposed.

Keywords — *Multi-primary display, metamerism, color, Binet-Cauchy theorem.*

1 Introduction

New display technologies that use more than three primaries have been introduced into the marketplace by several manufacturers.[1] These displays can have a wider color gamut than a conventional three-primary display whose gamut is restricted to a triangle in chromaticity space. Additional motivations for more than three primaries include better use of the photometric power of the light source and improved tone-scale rendering of reconstructed images on the display. A problem with these displays is that they increase the possibilities for on-screen metamerism, whereby observers may disagree on the equivalence of pairs of colors on the screen. The problem arises because, if a display has more than three primaries, many possible activations of the primaries exist that produce a given tristimulus specification. Existing displays restrict the activations so as to remove the ambiguity for, say, the 1931 CIE Standard Observer, but the non-uniqueness reasserts itself when the observer is changed through lens ageing or artificial lenses. This paper will characterize on-screen metamerism and provide design rules to mitigate it.

The use of more than three primaries may incur benefits in several niches of display technology. Today, most thin flat-panel active-matrix displays and projector displays employ a single broadband white light source, *i.e.*, a compact arc lamp source or a fluorescent backlight. The light source is modulated by amplitude-modulated spectral bandpass filters to produce discrete levels of several primary colors. In a liquid-crystal flat screen this is accomplished by combining a spatially distributed liquid-crystal amplitude-modulating light valve with a discrete color filter to define a sub-pixel primary color distributed uniformly over the display surface area. Flat-panel direct-view display technology relies on the eye's spatial blur to mix discrete primary color channels. Single-chip projection displays, *e.g.*, display devices that use a single spatial light modulation channel, modulate light in several ways that include liquid-crystal light valves, micro-mirrors, and diffraction. Here the mixing of the primaries is enabled by the eye's slow temporal response or persistence. Such a display device always uses a broad-spectrum arc lamp as the light source.

Solid-state light sources that are more efficient in converting electrons into photons in the visible spectrum are being developed for backlights and projector light sources. These future light sources are capable of temporal modulation that can be phased with spatially distributed amplitude modulation to reconstruct imagery. This paper addresses the problem of metamerism on current display devices and on future devices that will employ these new light sources.

All devices that employ a spatial or temporal multiplex, broad-spectrum light sources and three-color primaries require an engineering trade-off between color saturation and brightness. The introduction of more than three primaries gives the display designer additional latitude to generate saturated colors and yet still render high-contrast edges within the image. For example, the introduction of a white primary in a quad sub-pixel spatial arrangement, where each RGB triad pixel element is replaced with a RGBW quad, can increase the peak brightness of the display substantially and still use the full photometric power in the light source. The introduction of a white primary in a time-sequential projection display will also produce a brighter display with better color saturation. Since saturated colors tend to be dark, the addition of a white primary can substantially improve image quality, especially for natural images.

Quantization of the tone scale in digital electronic displays can produce an unwanted banding artifact in image regions that contain shallow intensity gradients. This artifact reduces image quality and is highly objectionable. It is produced because the intensity spacing in some regions of the tone scale is simply too big. The use of multiple primaries can help the system designer defeat this quantization artifact. Introducing additional color primaries in either spatial or temporal multiplexed systems necessarily reduces the

M. H. Brill is with Datacolor/Color Vision, 5 Princess Road, Lawrenceville, NJ 08648; telephone 609/895-7432, e-mail: mbrill@datacolor.com.
J. Larimer is with NASA Ames Research Center, Moffett Field, CA 90535-1000; e-mail: jlarimer@mail.arc.nasa.gov.

peak brightness of all the primaries. The choice of chromaticities and luminance ratios for the primaries gives the display designer more degrees of freedom to reduce the tone-scale step size between levels and thereby improve image quality by eliminating banding artifacts. However, an additional primary introduces the potential for metamerism. We address this problem and provide design rules for avoiding it.

2 On-screen metamerism: Some terminology

The CIE Standard Observer maps every possible screen color made by any combination of primaries into a tristimulus vector (each vector comprising three tristimulus values). For the standard observer, identical tristimulus vectors correspond to color matches. For every observer, screen patches that have the same composition of primaries are isomeric – *i.e.,* they have physically identical power spectra, and therefore they are also matched in color and brightness. This implies that they share the same Standard Observer tristimulus values. However, two lights that map to the same Standard Observer tristimulus values need not be isomeric. These lights, called metamers, are physically different lights that appear to be matched in color and brightness. Metameric equivalence is unique to the individual, so metamers defined by Standard Observer tristimulus values may not be metamers for real observers. The reason for this is that real observers differ from the CIE Standard Observer, mainly by the interposition of a non-neutral filter characterized by the individual's unique coloration in optical media and retinal pigmentation. Moreover, these coloration differences continue to change throughout life due to lens aging, artificial lenses, age-related changes in retinal pigmentation, or even wearing sunglasses.

On-screen metamerism (matching of spectrally different colors on the same display) does not arise for electronic displays that use three-color primaries. For such displays, all screen patches that are expected to appear identical are in practice isomers and therefore appear identical to all observers. Any set of three primaries employed in display applications are volume-filling primaries. That means the corresponding region of the Standard Observer's color space, the display's palette in CIE color space, has dimension three. This also implies that, in a three-color system, no two primaries can be admixed in any proportions to color-match to any amplitude (other than nil) of the third primary.

Three primaries that are volume filling to the Standard Observer may not be volume-filling for a trichromatic observer related to the Standard Observer by an interposed filter. To ensure volume-filling primaries for a filtered Standard Observer implies a constraint on the filter and a constraint on the spectra of the primaries. The primaries must satisfy a criterion to be developed in Section 4 of this article. The filter must be nonsingular, by which we mean that it does not completely exclude any part of the visible spectrum.

When a display system has $N > 3$ color primaries, metamers may not be isomers, and on-screen metamerism must be avoided by design choices. Two non-isomeric screen patches produced by an N-primary display that map to the same Standard Observer tristimulus values will almost certainly be seen as different by real observers. For example, suppose two patches have primary activations $\langle a_1, a_2, \ldots, a_N \rangle$ and $\langle b_1, b_2, \ldots, b_N \rangle$ $(0 \leq a_i, b_i \leq 1)$, which map into the same CIE 1931 tristimulus values for the standard observer, but for which individuals will not agree on their sameness because of the unique filter that each person looks through. In that case, lights that are not isomers, but metamers relative to the CIE standard, may not match to individual real observers. Also, the standard observer could see two screen colors as being different, whereas a real observer might see them as being the same. Either case produces on-screen metamerism.

Eliminating on-screen metamerism in effect means controlling the dimensionality of the subset of spectra selected from the full palette of possible spectra that can be produced on an N-primary display. The selected subset of screen colors will map uniquely into the CIE Standard Observer color space, so knowing the tristimulus values for any screen color uniquely determines the associated N-tuple.

The issue of on-screen metamerism is adjacent to some related topics that are not addressed by the present article.

First of all, real color matches are not transitive in practice: A matching B and B matching C does not ensure that A matches C. Color matches have random errors associated with the instrumentation and the observer. It is possible that two colors that are quite distinct to the CIE standard observer will become mathematically distinct but indistinguishable when a filter is interposed. Shall we call this a match? In this article we define a match as the mathematical equivalence, unaffected by "just-noticeable-difference" error. Such equivalence could conceivably have an operational foundation such as described by Stanford professor emeritus Joseph B. Keller: Two matching lights A and B "really match" if any light C that matches A also matches B. (Keller spoke at the New Jersey Institute of Technology Nov. 14, 2001; see summary by M. Brill in the ISCC News No. 394, Nov–Dec 2001, p. 16.)

Secondly, the prescription advanced here for avoiding display-induced metamerism remains subject to the increased color confusions that arise from observer color blindness or any singular filter (*i.e.,* a filter that reduces the number of unique Standard Observer tristimulus outcomes possible on the display). That is to say, dichromatic and monochromatic color observers are characterized by the property that they confuse a larger number of physically different lights than do normal trichromats. No modification to any electronic display can remedy this condition.

Thirdly, we do not address off-screen metamerism. On-screen metamerism involves comparing pixels on the same screen and not comparing on-screen colors with off-screen reference colors. Thus, avoidance of on-screen metamerism is very different from the problem of using sufficient primaries to approximate the power spectrum of an arbitrary patch of light in the outside world. When a display renders a scene in a colorimetrically exact way, on-screen metamerism does not disappear, but mimics what would happen in the world.

Finally, we defer to future efforts the characterization of the most likely personal filters to be encountered, and also design of primaries based on color-perception effects of such filters that transcend metameric matching. We are encouraged that, despite significant individually unique optical filtering of the light imaged onto human retinas, there is broad general agreement with respect to color appearance and the use of language to describe it.

3 Displays with three primaries and with three primaries at-a-time

A three-primary display avoids on-screen metamerism because each in-gamut tristimulus vector is produced by exactly one RGB triplet, a simple matrix-inverse giving the rule of correspondence. For an N-primary display, one can avoid on-screen metamerism by activating only three primaries at a time, so long as the primaries are well behaved under action of an optical filter. The primaries, taken in triads, should also tessellate the volume of the display's color palette. To determine which three primaries to use, one can first choose a triangular tiling of the chromaticity gamut of all the primaries, then locate the target color in one of these triangles, and finally produce the color using the primaries that define that triangle.

Of course, the choice of micro-gamuts, *i.e.*, primary triads, for defining the partition will depend upon several issues besides on-screen metamerism; for example, the luminous efficiency of the display device and image quality attributes such as avoiding tone-scale banding artifacts. Incorporating such issues is a topic for further investigation.

4 Binet-Cauchy condition for spectra of primary triplets

What constitutes good behavior of a triangle of primaries? If the primaries are monochromatic, their chromaticities are fixed, so no filter can change the triangle in which a target chromaticity lies. Therefore, one can always produce an in-gamut target chromaticity uniquely by using only the primaries of the triangle in which it lies. There can be no metamerism in that case. But having fixed chromaticities is not necessary for primaries to exhibit good behavior. It is necessary only that no triangle of primaries undergoes right/left-handed reversal under filtering, *i.e.*, that each triangle of chromaticities is such that a filter will not make one of the vertices cross the line between the other two. Such reversal could cause one triangle to flop on top of another, and there would be more than one allowed way to produce a color on the display. That would cause on-screen metamerism.

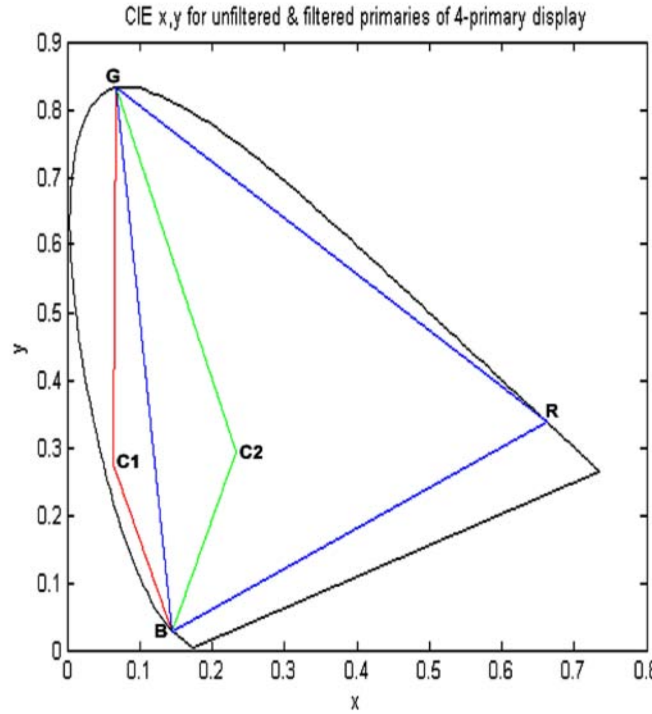

FIGURE 1 —Chromaticity space of a four-primary display, showing the migration of the C primary from C1 to C2 when a filter is introduced.

Figure 1 illustrates how such metamerism could happen. Imagine a four-primary display with monochromatic red, green, and blue (R, G, B) primaries at 610, 530, and 440 nm, and a cyan (C) primary that has 99% of its power at 500 nm and the rest at 610 nm. To the CIE observer, the chromaticity C1 of the C primary is almost but not quite on the spectrum locus. Let a filter be interposed that reduces the 500-nm contribution 20-fold, but transmits completely the long-wavelength end of the spectrum. The chromaticities of the R, G, B primaries are unaffected by the filter, but the C primary migrates across the line between G and B primaries, so it resides at chromaticity C2. Now, even if we agree to make colors only from triad BCG or triad BGR (depending on which triangle a target color inhabits), there are two ways in the filtered condition to make a slightly cyan color. If the target color is the C primary itself, we could use C or use a combination of R, G, and B. This dichotomy is on-screen metamerism: when the filter is taken away, the two ways of matching C become visually distinct.

The example above is somewhat extreme relative to today's displays because the primaries are narrow band. It serves nonetheless to illustrate the problem. In the future, when solid-state light sources start to be incorporated into display systems, this example will be strikingly similar to the

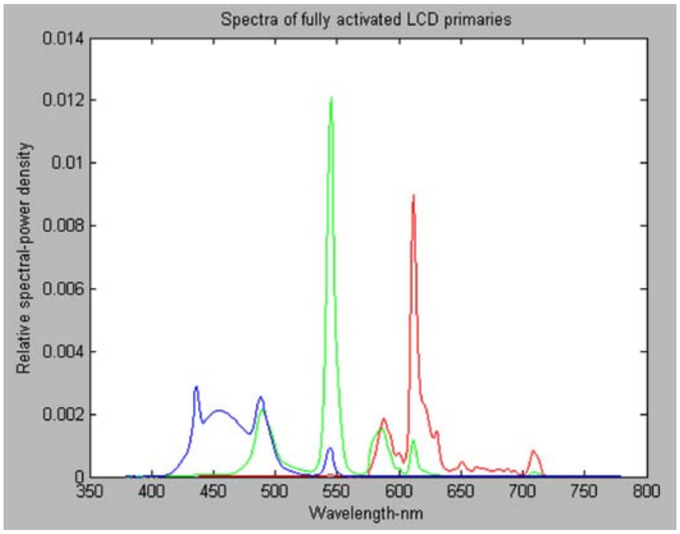

FIGURE 2 — Spectral power distributions of the primaries for an Apple® 17-in. LCD.

actual design choices awaiting the display design engineer. The non-crossing criterion is therefore important. To ensure it, one can impose a condition on the spectral power distributions* (SPDs) of the primaries, which is mathematically derivable using the Binet-Cauchy theorem.[2–6] Constrain the SPDs in wavelength λ of any three of the N primaries $P_i(\lambda)$, $P_j(\lambda)$, and $P_k(\lambda)$ so that the locus $[P_i(\lambda), P_j(\lambda)]/[P_i(\lambda) + P_j(\lambda) + P_k(\lambda)]$ is convex in wavelength λ. Primary spectra that inhabit compact non-overlapping wavelength domains will do the job, but others will too. A more detailed discussion of the Binet-Cauchy criterion appears in Appendix A.

It is worth asking at this point to what extent real primary spectra obey the Binet-Cauchy criterion. In Fig. 2, the three primaries of a high-end liquid-crystal display are shown. Their "chromaticity" plot is shown in Fig. 3, and one can see that the plot departs significantly from convexity. It is therefore possible to design a filter that will reverse the handedness of the chromaticities of the primaries. However, such a filter would have to exclude enough light at the high-emission wavelengths so that the nonconvexity (in parts of the spectrum with low emission power density) could effect the reversal. Such a personal filter would have little utility in terms of Darwinian notions of value and thus seems unlikely to occur naturally.

To use the Binet-Cauchy criterion in display design, one may need to accommodate primary spectra that are peaky, but in regions of the spectrum in which emitted power is low enough to make filter-induced chromaticity reversal unlikely. Accordingly, a "lenient" Binet-Cauchy criterion could be used that eliminates from the "spectrum locus" certain problematic wavelengths. For example, one could remove wavelengths whose power density in all three

*In all strictness, the spectrum of an illumination source is a density and transforms as such under domain changes such as wavelength-to-wavenumber, but the term "distribution" is always used instead of the more correct "density."

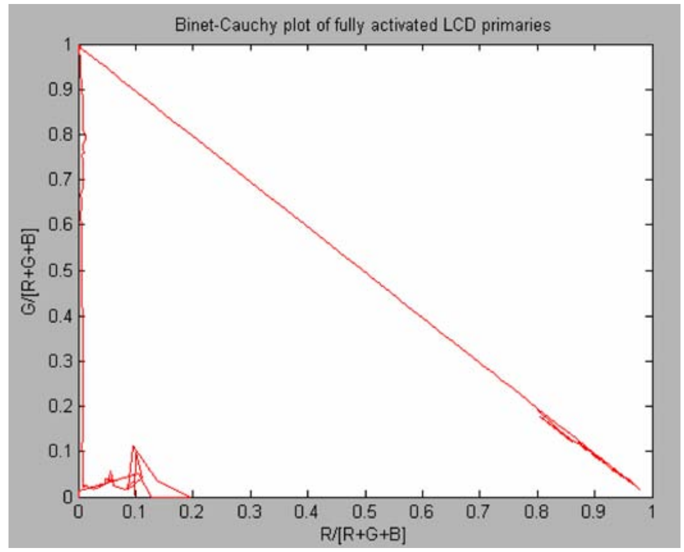

FIGURE 3 —Binet-Cauchy plot for the primaries in Fig. 2, from 400 to 700 nm. Here, R, G, B are spectral power densities as functions of λ.

primaries is below a criterion fraction of the sum of the emission peaks. In Fig. 4, the criterion fraction is chosen to be 0.01. It can be seen that the lenient criterion removes most – but not all – of the nonconvexity. But we note that if the primaries are narrow band spectrally, then narrow-band filtering can easily generate singular (*i.e.*, dimension-reducing) transforms on the tristimulus values either for the Standard Observer or for the tristimulus space defining each unique observer. In this case, enlarged equivalence classes of metamers are likely. A display design criterion might be to make the reconstructed colors as much like the reference colors as possible. Further study will be needed to refine the Binet-Cauchy criterion to a practical standard and to address the problem of the off-screen reference color.

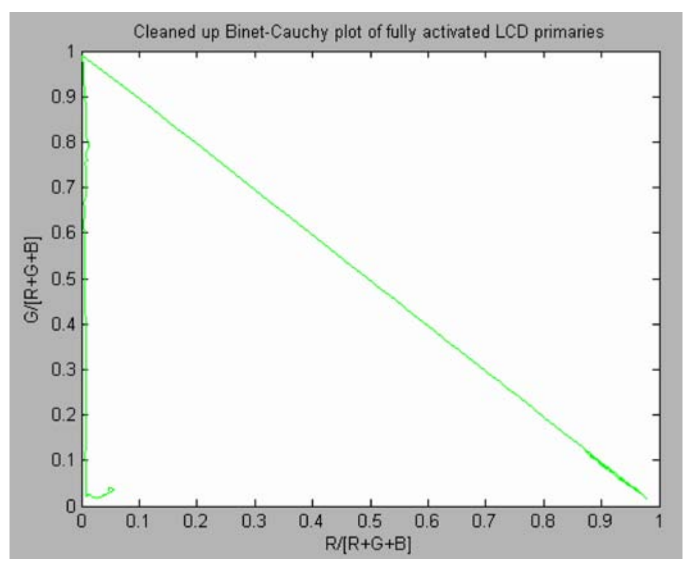

FIGURE 4 —Binet-Cauchy plot of Fig. 3, with threshold at 0.01 max (R + G + B) over λ. As in Fig. 3, R, G, B are spectral power densities as functions of λ.

5 Composite primaries from monochromatic simple primaries

The definition of on-screen metamerism implies that its avoidance *requires* use of only three primaries (additive degrees of freedom) at a time. At first this requirement seems to limit luminance for chromaticities that could otherwise be made from more than three of the N primaries. However, we can get more luminance by expanding the definition of "additive primary" to include *composite* primaries, fixed linear combinations of the N primaries.

Composite primaries defined relative to the Standard Observer's tristimulus values are column-vectors $\mathbf{R}$ for a target color and $\mathbf{R}_k$ (k = 1, ..., N) for fully activated primary k. (Note: For monochromatic primaries, the chromaticity of each primary is invariant to action of any filter over the eye, but not so for the target color, of course.) If a target color is a combination of three composite primaries, then $\mathbf{R}$ is a linear combination of three fixed linear combinations of the N simple primaries $\mathbf{R}_k$:

$$\mathbf{R} = \sum_{j=1}^{3} d_j(R) \sum_{k=1}^{N} A_{jk}\mathbf{R}_k. \tag{1}$$

Here, the $\mathbf{R}$-independent $3 \times N$ matrix $\mathbf{A}$ defines the composite primaries. These primaries are linearly combined to match $\mathbf{R}$, using weights $d_j(R)$ that depend on $\mathbf{R}$.

The eye integrates radiant power imaged on the retina spectrally, spatially, and temporally, so if the simple primaries are monochromatic, composite primaries derived from them allow a greater luminance of chromaticities near white than would be obtained from the simple primaries alone.

There are many possible ways to create a three-at-a-time system for using composite primaries of a given N-primary display. First, designate a set of composite primaries and tile chromaticity space with triangles of them, out to the limits of the multi-primary gamut. For any target chromaticity, render the color using only the three primaries (simple or composite) in the uniquely defined triangle in which the target color resides (in unfiltered state). In the process, be sure the composite primaries (whose chromaticity depends on the filter) cannot reverse the handedness of any chromaticity triangles; *i.e.,* make sure that no interposed filter can cause a composite primary to cross the line between two other of the primaries, either simple or composite.

For example, let the monochromatic primaries be B, C, G, Y, R, in obvious spectral order. Define doublet composite primaries as tristimulus sums of spectrally adjacent primaries: B + C, C + G, G + Y, Y + R, R + B. Filtering causes a doublet primary to move only along the line between its constituent simple-primary chromaticities – and these simple-primary chromaticities are filter-invariant. Therefore, no triangle comprising any three of these primaries (simple or composite) will undergo any chromaticity reversal by application of a filter. Finally, assign a single white composite primary S = B + C + G + Y + R, which is guaranteed to lie within the pentagon of the doublet primaries, and hence within the pentagon of the simple primaries. Then no triangle among any of the primaries (singlet, doublet, or pentuplet) will reverse with filtering. Accordingly, denote the target-color triangular areas (B, B + C, R + B), (C, B + C, C + G), (G, G + C, G + Y), (Y, Y + G, Y + R), (R, R + Y, R + B), (S, R + B, B + C), (S, B + C, C + G), (S, C + G, G + Y), (S, G + Y, Y + R), and (S, Y + R, R + B). This is an affinely robust sort of "Fuller dome" anchored in the monochromatic primaries. (Of course, there is still a luminance price to pay for intermediate-chroma colors, but this price will be bearable so long as N is not too large.)

It should be noted that the chromaticities of the composite primaries are not filter-invariant, even if the individual primaries have filter-invariant chromaticities. Therefore, only the Binet-Cauchy criterion will keep a color gamut of more than three composite primaries from folding on itself (hence, becoming metameric) as one applies a filter.

6 Auxiliary conditions and composite primaries

This section deals with an alternative mathematical picture of composite primaries. We include it only to prevent the misconception that the picture describes something other than composite primaries. Accordingly, the section can be skipped with no practical implementation consequences.

For $N > 3$ monochromatic primaries, one can avoid metamerism by imposing $N - 3$ auxiliary linear conditions along with the tristimulus matching conditions. Define $\mathbf{R}$ and $\mathbf{R}_k$ as before. Define an N-vector $\mathbf{b}$ of commanded gains imposed for the N primaries to match $\mathbf{R}$ to the unfiltered eye, so that

$$\mathbf{R} = b_1\mathbf{R}_1 + b_2\mathbf{R}_2 + \ldots + b_n\mathbf{R}_N. \tag{2}$$

Impose the following auxiliary conditions on $\mathbf{b}$ so as to be able to solve uniquely for $\mathbf{b}$:

$$c_i = a_{i1}b_1 + a_{i2}b_2 + \ldots + a_{in}b_N, \tag{3}$$

where i = 1, ..., $N - 3$, c_i are selected constants, and $\mathbf{R}_k$ are the tristimulus vectors of the fully-activated primaries – whose chromaticities are invariant to filter action. Of course, the black point $\mathbf{R} = 0$ must be allowed, so zero values for all the b_k must be allowed, and that forces all the c_i to be 0. Care must be taken so that the auxiliary condition vectors (a_{i1}, a_{i2} ... a_{in}) are linearly independent of the other row vectors defined by ($\mathbf{R}_1$, ..., $\mathbf{R}_N$). Denote by $\mathbf{M}$ the $N \times N$ matrix whose kth column comprises (in order) $\mathbf{R}_k$ and a_{ik}, and denote by $\mathbf{r}$ the column N-vector comprising (in order) $\mathbf{R}$ and ($N - 3$) entries 0. Then we have $\mathbf{r} = \mathbf{M}\,\mathbf{b}$, and hence can impose the gains $\mathbf{b} = \mathbf{M}^{-1}\,\mathbf{r}$, predicated on the unfiltered eye.

It can be shown (see Appendix B) that using auxiliary conditions is equivalent to using composite primaries. However, the composite-primary picture is more instructive than the auxiliary-condition picture. For one thing, the use of three *simple* primaries emerges as a clear sub-case. Also,

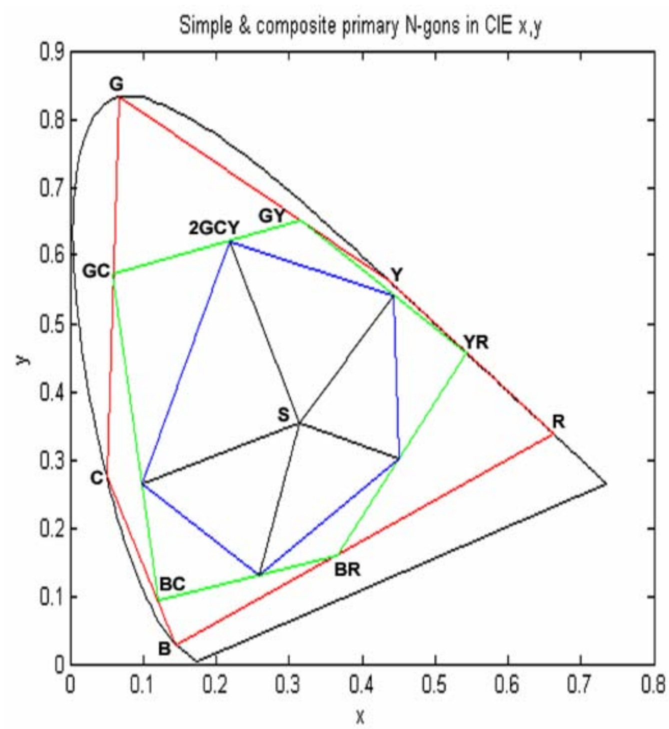

FIGURE 5 —Color space for a display with five monochromatic primaries. Here, the chromaticity GC is the chromaticity of the tristimulus sum of G and C, abbreviated as GC = G + C. Similarly, 2GCY = G + C + G + Y, *S* = B + C + G + Y + R = BC + BR + YR + GY + GC, *etc*.

some coefficient constraints are implied by the fact that only nonnegative scales can be applied to each composite primary, and only nonnegative relative scales among the simple primaries can comprise a viable composite primary. Thus each composite-primary chromaticity must lie within the gamut of the simple primaries that comprise it.

7 General use of composite primaries

The use of composite primaries in Section 5 can be extended further, in two ways:

(1) The simple primaries need not be monochromatic so long as their spectral power densities satisfy the Binet-Cauchy criterion for all possible triangles (triads of simple primaries) created from the *N*-gon of primaries. Also the *N*-gon of primaries should be convex.

(2) The procedure for using composite primaries need not stop after one iteration. One can develop an arbitrarily large sequence of concentric *N*-gons, each successive *N*-gon's vertices lying on the respective sides of the previous (larger) *N*-gon. This process automatically defines a set of *N* triangles between each successive pair of *N*-gons, and those triangles can never invert with scaling of the simple primaries. When one chooses to end the process, the innermost *N*-gon vertices are connected to the chromaticity *S* that is the sum of the outermost *N*-gon tristimuli. That last process creates *N* more triangles, which are also not invertible by scaling of the simple primaries. This process works because the point *S* is the tristimulus sum of the vertices of any one of the *N*-gons, and hence filtering will not cause *S* to cross any side of any *N*-gon.

To illustrate (2) above, add one more step to the example in Section 5. The third-level composite primaries (comprising a pentagon) are defined by 2C + B + G, 2G + C + Y, 2Y + G + R, 2R + Y + B, and 2B + R + C. These primaries are sums of the second-level composite primaries defining the second-level pentagon. The method evolves new triangles of the form CB, (2B + R + C), (2C + G + B), and so forth. If one stopped generating pentagons at the third level, the third-level vertices would be joined to *S* to give triangles such as S, (2B + R + C), (2C + G + B). Proceeding through *M* iterations on the basic-primary *N*-gon produces MN triangles, none of which will reverse with filtering of the basic-primaries. Note that this bootstrap approach avoids the luminance-sagging problem of using only three simple primaries at a time. The only drawback is the multiplicity of triangles, which will impose a proportionate computational load.

8 Example of primary-set construction

The following algorithm implements the above considerations:

Step 1. Make sure that any three of the *N* primary spectra obey the Binet-Cauchy criterion (so the *N* do not tangle up when filtered). Also assure the convexity of the chromaticity-space *N*-gon of the primaries (to avoid defining overlapping triangles in steps 2–5 below).

Step 2. Starting with the original (parent) *N*-gon of the simple primaries, add the tristimulus vectors of the primaries from adjacent vertices of the *N*-gon to make second-level composite primaries (an *N*-gon inscribed in the parent *N*-gon). The triangles defined by the sides of the *N*-gons are primary triads, and a target color within one triangle should be rendered only by that triangle of primaries.

Step 3. Treat the second-level *N*-gon as the new parent, and generate the third-level *N*-gon. Define primary triangles between the second-level and third-level *N*-gons.

Step 4. Continue making triangles out of adjacent *N*-gons as long as desired. (Of course, one could have skipped steps 2 and 3 or step 3 if only one or two *N*-gons were needed.) Then connect the vertices of the innermost *N*-gon with the sum *S* of the parent primaries. That generates *N* more triangles.

Step 5. Given an in-gamut target (x,y,Y), find its triangle among the *NM* triangles generated in Steps 2–4, and then use the primaries at the vertices of that triangle to render the color on the *N*-primary display.

The above algorithm is illustrated by the chromaticity diagram in Fig. 5. The chromaticities of five simple monochromatic primaries B, C, G, Y, R are connected by the red pentagon. The monochromatic primaries are presumed to have the same power at full activation, and have wavelengths 440, 470, 500, 550, and 630 nm. Second-level prima-

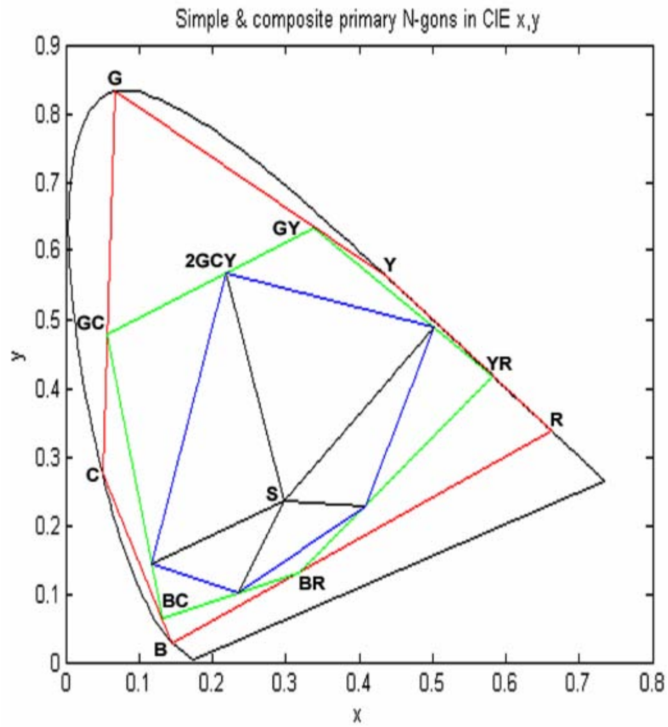

FIGURE 6 —Color space for the display of Fig. 5, but with filtered primaries.

ries BC (which is the tristimulus sum of B and C), GC, GY, YR, and BR are connected by the green pentagon. Third-level primaries (*e.g.*, 2GCY = 2G + C + Y) are connected by the blue pentagon. The sum of all the primaries at first, second, or third level, has the chromaticity of the white *S* at the center of the diagram.

Figure 6 illustrates the distortion of the five-primary system in Fig. 5 due to filter factors 1, 1/2, 1/4, 1/3, and 2/3 (from shortest to longest wavelength). Filtering leaves the red pentagon fixed, slides the green vertices along the red-pentagon sides, slides the blue vertices along the green pentagon sides, and moves the white point within the blue pentagon. These actions do not reverse any of the depicted triangles. Note that even with such severe filtering, the latticework of the primaries (triangle vertices) remains undisturbed in its order.

9 Outlook

Practical realities may mitigate the direct use of the algorithm described in Section 8. For example, real primaries might violate the strict Binet-Cauchy criterion, in which case gentler versions of that criterion will be needed.

If the chromaticity polygon of the simple primaries is not convex (as required in Step 1 of the algorithm of Section 8), then the algorithm must be somewhat modified. For example, if a white LED provides one of the simple primaries (*W*) and the other simple primaries form a convex $(N - 1)$-gon that contains *W*, that $(N - 1)$-gon is treated as the first polygon in the algorithm. The Binet-Cauchy test is then applied not only to the simple primaries, but also to all triplets involving the composite primaries and *W*. If *W* lies inside the innermost *N*-gon and passes the Binet-Caushy test with all doublets of composite primaries, then *W* replaces *S* as the display white.

It should also be remembered that real displays rarely are completely additive in their primaries: CRTs tend to have weaker-than-additive white because of beam-current loading and the use of limited output power supplies, and LCDs do not achieve neutral blacks due to the dispersive nature of the light valve. In a normally white LCD, for example, the chromaticity of the black level tends to be bluish due to incomplete extinction of light passing through the blue sub-pixel elements. The cell gap in these displays is set to extinguish the sub-pixel with the greatest luminosity, which is always the green primary because its dominant wavelength is closest to the photopic V_λ maximum. The primaries of display devices that are not based on liquid crystal light valves may also depart somewhat from chromaticity invariance with respect to activation level. All these non-ideal behaviors may cause overlapping between regions driven by two neighboring triads of primaries. If the inversions occur at low screen luminance (as most often is the case), their expected conspicuity would be low.

In summary, the challenges that remain in display design so as to mitigate on-screen metamerism seem manageable.

References

1 S Roth, I Ben-David, M Ben-Chorin, D Eliav, and O Ben-David, "Wide gamut, high brightness multiple primaries single panel projection displays," *SID Symposium Digest* **34**, 118–121 (2003).

2 H W Turnbull, *The Theory of Determinants, Matrices, and Invariants* (London: Blackie, 1945), Chapter V.

3 C C MacDuffee, *The Theory of Matrices* (New York: Chelsea, 1946), p. 9 (Theorem 7.9).

4 M H Brill, "Decomposition of Cohen's matrix R into simpler color invariants," *Am J Psychol* **98**, 625–634 (1985).

5 M H Brill and H Hemmendinger, "Illuminant dependence of object-color ordering," *Die Farbe* **32/33**, 35–42 (1985/86).

6 M H Brill, "Color-science applications of the Binet-Cauchy Theorem," *Color Research and Application* **27**, 310–315 (2002).

7 G West and M H Brill, "Conditions under which Schrödinger object colors are optimal," *J Opt Soc Am* **73**, 1223–1225 (1983).

Appendix A. Binet-Cauchy Criterion

Let $P_i(\lambda)$ be the spectral power distribution of primary *i* in a display. We will deal with three primaries at a time, so without loss of generality $i = 1, 2, 3$. Let $f(\lambda)$ be a filter transmittance, and let $q_j(\lambda)$ be a set of CIE color-matching functions. Then the tristimulus values Q_{ji} produced by primaries P_i seen through filter *f* are given by

$$Q_{ji} = \int f(\lambda) q_j(\lambda) P_i(\lambda)\, d\lambda, \qquad \text{(A-1)}$$

It is possible to write Eq. (A-1) as

$$\mathbf{Q} = \int f(\lambda) \mathbf{q}(\lambda) \mathbf{P}^{\mathrm{T}}(\lambda)\, d\lambda, \qquad \text{(A-2)}$$

where $\mathbf{Q}$ is the 3×3 matrix with the components Q_{ji}, $\mathbf{q}(\lambda)$ is a column-vector-valued function of λ comprising $q_j(\lambda)$, and $\mathbf{P}(\lambda)$ is a column-vector-valued function of λ comprising $P_i(\lambda)$, and superscript T denotes matrix transposition.

Expressing the wavelength as a finite index ($k = 1,\ldots,N$), and the integral as a Riemann sum from 1 to N, converts Eq. (A-2) to

$$\mathbf{Q} = \mathbf{A}^{\mathrm{T}}\mathbf{B}, \tag{A-3}$$

where $A_{kj} = f(k)q_j(k)$ and $B_{ki} = P_i(k)\, d\lambda$.

Now, the handedness of the chromaticity ordering of P_1, P_2, P_3 changes (*i.e.,* the triangle flips) when the algebraic sign on the determinant $\det(\mathbf{Q})$ changes. Therefore, if the sign on $\det(\mathbf{Q})$ does not change with f, then the ordering is filter-invariant. To show that $\mathrm{sgn}[\det(\mathbf{Q})]$ does not change with f, we invoke a restricted case of the Binet-Cauchy theorem[1–3]:

$$\det(\mathbf{Q}) = \sum_{k,m,n} \det[\mathbf{A}(k,m,n)]\det[\mathbf{B}(k,m,n)], \tag{A-4}$$

where $\mathbf{A}(k,m,n)$ is the 3×3 block of $\mathbf{A}$ that comprises columns k, m, n, and $1 \le k < m < n \le N$. Each determinant $\det[\mathbf{A}(k,m,n)]$ evaluates to $f(k)\ f(m)\ f(n)\ \det[\mathbf{q}(k,m,n)]$, whose sign is the same as that of $\det[\mathbf{q}(k,m,n)]$. {**Note:** Here, $\det[\mathbf{q}(k,m,n)]$ is an abbreviation for $\det[\mathbf{q}(k), \mathbf{q}(m), \mathbf{q}(n)]$.} Because the spectrum locus is convex in wavelength, the sign of $\det[\mathbf{q}(k,m,n)]$ is independent of the choice of k, m, and n. (The convexity of the CIE spectrum locus is responsible for the optimal-color solid having the familiar stop-band and pass-band 1-0 reflectances noted by Ostwald, Schroedinger, and MacAdam.[7]) Therefore, the sign of $\det(\mathbf{Q})$ will be observer filter-invariant if the three primary spectra have the following *convexity property*: The sign on $\det[\mathbf{B}(k,m,n)]$ does not depend on the choice of k, m, n, for $k < m < n$.

The above statement of the convexity property is equivalent to the following: If the primary spectra $P_1(\lambda)$, $P_2(\lambda)$, $P_3(\lambda)$ are considered analogous to color-matching functions, they create a "spectrum locus" that is convex and well ordered in wavelength. Hence, the 2-space point with coordinates $c_1 = P_1(\lambda)/[P_1(\lambda) + P_2(\lambda) + P_3(\lambda)]$, $c_2 = P_2(\lambda)/[P_1(\lambda) + P_2(\lambda) + P_3(\lambda)]$ traces out a convex curve in parameter λ. Primary spectra that trace out such a curve will be said to satisfy the Binet-Cauchy criterion. (Note: if the curve retraces itself or doubles back on itself, it is not convex.)

Appendix B. Proof that auxiliary conditions are equivalent to composite primaries

Combining Eqs. (1) and (2) implies

$$b_k = \sum_{j=1}^{3} d_j(R)A_{jk}. \tag{B-1}$$

Substituting Eq. (B-1) into Eq. (3) (with $c_i = 0$, of course) and eliminating b_k yields

$$0 = \sum_{k=1}^{N} a_{ik} \sum_{j=1}^{3} d_j(R)A_{jk}. \tag{B-2}$$

Because $d_j(R)$ depends on $\mathbf{R}$, each coefficient of $d_j(R)$ must individually be zero, hence

$$0 = \sum_{k=1}^{N} a_{ik}A_{jk} \tag{B-3}$$

for $i = 1,\ldots,N-3$ and $j = 1, 2, 3$. This means that the row vectors of $\mathbf{A}$ are orthogonal to the row vectors of **[a]**. The isomorphism is thereby shown: the composite-primary picture deals with the three-dimensional subspace of the primaries, whereas the auxiliary-condition picture deals with the $(N-3)$-dimensional orthogonal complement space of co-dimension 3.

Michael H. Brill is the Principal Color Scientist at Datacolor/ColorVision in Lawrenceville, New Jersey. Since obtaining his Ph.D. in physics at Syracuse University, he has carried out extensive theoretical research in color in human and computer vision, in geometric/photometric invariance, and in physics-based vision. He is co-inventor of the Emmy-Award-Winning Sarnoff vision model. He is also a Past President of the Inter-Society Color Council, and author of more than 80 refereed technical publications, 10 U.S. patents, numerous national standards, and a SID test pattern. He has chaired or co-chaired five conferences on color technology, vision, and digital display. In addition, he obtained the 1996 Macbeth Award from the ISCC for his work on color constancy. Dr. Brill is a member of the Editorial Board of Color Research and Application and chairs CIE Technical Committee TC1-56: Improved Color Matching Functions.

James Larimer received the Ph.D. degree from Purdue University and was an NIH postdoctoral Fellow at the Human Performance Center of the University of Michigan, Ann Arbor. His doctoral and post-doctoral work specialized in quantitative methods, statistics, and human sensory physiology and psychophysics. For 15 years after leaving the University of Michigan he was a Professor, and later Department Chairman, at Temple University. He taught sensory physiology and perception, statistics and experimental design. In the mid-80's, he was the Director of the National Science Foundation's Sensory Physiology and Perception Program. In 1987 he joined NASA's Ames Research Center as a Senior Scientist.