Chapter 12

Psychophysics of Postreceptoral Color Codes

Contents

12.1 Psychophysical paradigms that reveal postreceptoral codes . . . . . . . 298
12.2 Paradigm 1: Summation-near-threshold . . . . . . . . . . . . . . . . . . . 298
12.3 Summation experiments in color space . . . . . . . . . . . . . . . . . . . 301
12.4 Paradigm 2: Adaptation-near-threshold . . . . . . . . . . . . . . . . . . . 309
12.5 Measuring flicker adaptation in color space . . . . . . . . . . . . . . . . . 311
12.6 Summary of evidence for the Boynton code . . . . . . . . . . . . . . . . . 316
12.7 Photometry revisited . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 316
12.8 Summary . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 320

In Chapter 11 we introduced the concept of postreceptoral processing in color vision – the idea that the initial signals from the L, M, and S cones are recombined early in the visual system into a new color code. We illustrated the concept of recodings at the theoretical level by discussing linear three-channel color models, and the color spaces in which they can be represented, with a change of color code being represented by a change in the axes of a three-dimensional color space.

In the present chapter, we follow through these ideas in the psychophysical realm. The question is, are there ways in which system properties determined from psychophysical experiments can provide any evidence about postreceptoral neural codes? At first glance this challenge seems an impossible one. The visual system might seem like an impenetrable black box, and there might seem to be no way to deduce the wiring without opening the box.

In fact, several psychophysical paradigms have been developed and used, within specific theoretical contexts, to glean information concerning postreceptoral code changes. In the present chapter we will introduce two of these paradigms, and show how they play out in the context of color vision. Later (Chapter 15), we will use the same two paradigms to explore some surprising characteristics of postreceptoral processing in spatial vision.
12.1 Psychophysical paradigms that reveal postreceptoral codes

In her influential book, *Visual Pattern Analyzers*, Norma Graham (1989) provided an in-depth analysis of four psychophysical paradigms that provide information about postreceptoral visual processing. We will discuss two of these paradigms: *summation-near-threshold* and *adaptation-near-threshold*. In each case, after a brief general discussion, we will show how the paradigm is applied in the context of color vision.

12.2 Paradigm 1: Summation-near-threshold

12.2.1 The summation square: Summation, subtraction, or independence?

Suppose you have measured detection thresholds for two stimuli, A and B (for example, A and B might be test spots of 540 and 650 nm respectively). Suppose you now present the two stimuli simultaneously, superimposed at the same location. We will call A and B the *component stimuli*, and call the superposition of A and B the *compound stimulus*. It is denoted A ⊕ B. And suppose you are interested in whether or not the signals S_A and S_B, set up by stimuli A and B, interact within the visual system, and if so, in what way.

A rather dramatic way to express the possible outcomes of such an experiment is in a *summation square*, as shown in Figure 12.1. In the summation square, the x and y axes represent the intensities of the stimuli A and B. The units on both axes are normalized to detection thresholds, so that 1 unit on the x axis denotes the detection threshold for A, and 1 unit on the y axis denotes the detection threshold for B. With this normalization, the thresholds for A alone and B alone are represented by the solid circles at the points 0,0 and 1,0 respectively.

Now, imagine superimposing components A and B, and measuring thresholds for the compound stimulus. Let’s set the normalized intensities of A and B to a fixed ratio – say 1:1. As we vary the intensity of the compound, the stimulus will vary along a ray out from the origin, as shown in Figure A. The subject’s task is to vary the intensity of the compound to find its detection threshold. The experiment is repeated with different fixed ratios between A and B, yielding a set of detection thresholds, represented by the black dots. A line connecting these threshold values can be called a *detection contour*. The question is, what pattern of threshold values – what detection contour – will be traced out for the compound stimulus?

The three solid lines in Figure 12.1B represent detection contours predicted from three very different rules for recombining the signals from A and B. First, if the signals are processed by a channel that exhibits linear *summation*, the threshold for the compound will be reached whenever the sum of the signals, S_A + S_B, reaches a value of 1. The higher the intensity of A, the lower the intensity of B that will be needed for the compound to be at threshold. The detection contour predicted by linear summation model is represented by the solid diagonal line that runs from (0, 0) to (1, 1), and represents the equation S_A + S_B = 1.

Second, what would happen if the signals from the two stimuli were combined with opposite sign, or *subtracted* from each other? The two signals should cancel. The higher the intensity of A, the higher the intensity of B that would be needed for detecting the compound stimulus. Detection contours predicted from a *subtractive* (or *opponent*, or *antagonistic*, or *cancellative*) model are represented by the two parallel diagonal lines that originate from the points at (1, 0) and (0, 1), and rise outside the summation square with slopes of 1. (The two lines actually represent
12.2. PARADIGM 1: SUMMATION-NEAR-THRESHOLD

Figure 12.1: Summation squares. The axes represent the intensities of two stimulus components, A and B, normalized so that the detection threshold for each component is set to the value of 1. A. Detection thresholds are measured for various compounds (mixture ratios) of A and B. The detection threshold for each A:B ratio is marked with a dot. The line joining the threshold values is called a detection contour. B. Detection contours predicted from three major recombination rules: summation, subtraction, and independence. C. A set of detection contours predicted from the rule $S_A^k + S_B^k = 1$, for various values of k. D. Data from a summation-near-threshold experiment using 540 and 650 nm stimulus components. [A,B: DT. C: modified from Graham (1989, Fig. 4.9, p. 173). D: modified from Thornton and Pugh (1983, Fig. 2, p. 192).]
predictions from two different subtractive rules, $S_A - S_B = 1$ and $S_B - S_A = 1$.) [Figure out the predictions for weighted differences; e.g., 2L - M, 3L - M, etc.]

And third, what would happen if the two stimuli were detected independently in a system with two independent detection channels? Neither stimulus should affect the threshold for the other. The compound stimulus would be detected whenever either $S_A$ or $S_B$ reaches its individual detection threshold value of 1. So the detection contour predicted by a two-channel, independent detection model is a square, defined by the line segments from (0, 1) to (1, 1) and from (1, 0) to (1, 1).

In addition to summation, subtraction, and independence, there are options that fall in between. A general class of models used to predict outcomes that fall between the summation and independence contours posits that the signals $S_A$ and $S_B$ will be raised to the exponent $k$ before being combined, so that the signal from the compound stimulus is $S_A^k + S_B^k$. As shown in Figure 12.1C, $k = 1$ yields the prediction of linear summation, already discussed; $k = 2$ predicts a quarter circle; and larger exponents yield contours that bow out to approach the independence prediction more and more closely. These mathematical models are not closely tied to physiological models, but they are often used to describe the results of summation-near-threshold experiments.

An important detail used in interpreting summation-near-threshold data is an effect called probability summation. Assume we are doing a two-alternative forced-choice experiment, for detecting component A at threshold the subject’s proportion correct is 0.75. Assume a similarly threshold for component B. By analogy, imagine we are tossing two biased coins, each with the probability of a head set at 0.75. Assuming statistical independence, the probability of detecting the compound corresponds to the probability of detecting either or both of the components – getting at least one head. This probability is 1 minus the probability of getting two tails, or $1 - (.25)(.25) = .94$. That is, when the probability of detecting each component is 0.75, the probability of detecting the compound would be greater than 0.75, and to find the threshold for the compound the subject would need to turn down the intensity of the compound below the intensity needed for either component alone. In graphical terms, the effect of probability summation is to round off the corner of the independence prediction, and make it look like a summation contour with a high value of $k^2$.

Finally, if the two gratings facilitated each others’ detection – a low intensity of stimulus A makes stimulus B much more detectable, and vice versa – then the data would bow inward from the linear summation line. And subtractive models with various weightings of different cone signals yield diagonals of various slopes outside the square.

In summary, summation-near-threshold experiments are a powerful paradigm for exploring the recombination of neural signals within the visual system. A major advantage is that, as illustrated in Figure 12.1B, the predictions from the three major recombination rules – summation, subtraction, and independence – differ enough to be distinguished readily with experimental data. Thus, this paradigm could potentially give us an important tool for attacking the question of how signals from the L, M and S cones are recoded into a new postreceptoral code.

Notice that models that posit detection of the signals generated by the stimuli A and B by two independent channels, predict detection contours aligned with the original axes of the summation.

---

1 The exponent $k$ is known as the Minkowski metric. The case of $k = 1$ is linear summation, as described above. The case of $k = 2$ is sometimes called Euclidian summation because $x^2 + y^2 = c$ is the formula for a circle, so $k = 2$ predicts a quarter circle.

2 There are two uses of independence in this paragraph that are distinct. The independence prediction refers to the threshold determined by one component stimuli is independent of any subthreshold value of the other component across conditions. The statistical independence assumed for probability summation is that for any one condition, the probability of detecting one component doesn’t depend on the probability of detecting the other component.
square. In contrast, models that posit interactions – addition or subtraction of the signals from A and B – predict two different kinds of diagonal detection contours. Thus, by comparing detection thresholds for the two component stimuli vs. the compound, we should be able to determine the combination rules for the signals from the two components. This pattern of predictions – contours aligned with the axes signalling independence, and contours diagonal to the axes signalling interactions – will be with us throughout our treatment of summation squares.

12.2.2 Color codes: Evidence from a summation-near-threshold experiment

In the context of color vision, an interesting example of data taken with the summation-near-threshold paradigm is a study by Thornton and Pugh (1983). In this experiment, the authors first determined detection thresholds for two component stimuli – 650 and 540 nm test spots. They then measured detection thresholds for compounds of the two. The test stimuli were spatially blurred and ramped gradually on and off, and all thresholds were measured against a high intensity, 580 nm background.

Thornton and Pugh’s data are shown in a summation square format in Figure 12.1D. As before, the axes of the summation square are normalized to the detection thresholds for the two component test lights. The detection contour for the compound stimulus clearly shows subtraction – adding 540 nm light makes the 650 nm stimulus much harder to see than it was when it was presented alone, and vice versa. Because only wavelengths above 540 nm were used, and S cones are virtually unresponsive in this spectral region, we can guess that the subtractive interaction is between L- and M-cone-initiated signals. This result allows one to reject a model that summates the signals from 540 and 650 nm. Instead, the results point to a channel that subtracts the signals initiated from these two lights.

12.3 Summation experiments in color space

12.3.1 Detection contours in L, M cone input space

A more sophisticated approach would be to work in a cone contrast space. Let’s go to a two-dimensional L, M cone contrast space. Recall from the prior chapter that a cone space represents the excitation of each class of cone and is refined in a cone contrast space with respect to a background light that defines the origin and the axis values that are contrasts which depend on the background light. Rather than representing two arbitrary stimuli A and B, the x and y axes will now represent stimuli custom designed to isolate the L and M cones respectively. Therefore – and here is the main point – the different possible threshold contours will correspond, not to different rules for combining arbitrary signals $S_A$ and $S_B$, but to different rules for combining L-cone- and M-cone-initiated signals. The predicted detection contours are shown in Figure 12.2A. Remarkably,
12.3.2 Detection contours in L, M cone contrast space

Now let’s take one more step up in sophistication. Rather than just measuring thresholds for increments of light as Thornton and Pugh did, let’s measure modulation thresholds. That is, the stimulus will be a sinusoidal modulation through the origin along some particular axis of L, M cone contrast space, as shown in Figure 12.2B. For example, if the stimulus is a temporal modulation, the stimulus will modulate smoothly from, say, a high to a low intensity “white” light, or a high intensity “red” to a low intensity “green”, or between isoluminant “red” and “green”. The subject’s task is to vary the modulation depth along the chosen axis until the modulation is just barely visible; that is, to measure a modulation or contrast threshold. The experiment is repeated on many axes through the origin, to generate a set of contrast thresholds for modulations in many directions in L, M cone contrast space. By convention, each contrast threshold is recorded with a pair of dots, one on each side of the origin on the axis in question.

What shape of detection contour should be traced out by contrast thresholds in L,M cone contrast space? Different combination rules predict different contours, as shown in Figure 12.3. For a channel that sums L and M cone signals – an L + M channel – the predicted detection contour is given by points that fit the equation L + M = 1. This prediction is a pair of parallel lines passing through the points (1, 0) and (0, 1) with negative slope, as shown by the dashed lines in Figure 12.3A. For a channel that subtracts L vs. M signals – an L - M channel – the prediction is given by points that fit the equation L - M = 1. The prediction is a pair of parallel lines passing through the points (1, 0) and (0, 1) with positive slope, as shown by the dashed lines in Figure 12.3B. But if (contrary to the Boynton code) the L and M cone signals do not interact, but rather feed into separate, independent channels, the prediction is that the set of modulation thresholds should trace out a square, aligned with the L and M axes, as shown by the dashed lines in Figure 12.3C.

Moreover, these predictions can be extended to cover variations in the sensitivities of the L + M and L - M channels. That is, we expect that as we vary (say) the temporal frequency of the stimulus, the sensitivities of the different channels will vary. The less sensitive the L + M channel (say) is under the conditions tested, the farther apart will be the two parallel lines in Figure 12.3A. This effect is illustrated by the two dotted lines for the case in which the sensitivity of the L + M channel has been reduced by a factor of two, so that a signal of the value 2 is needed for detection. The same argument holds for the L - M channel (Figure 12.3B), and for independence (Figure 12.3C).

What if the system has both an L + M and an L - M channel? This case is shown in Figure 12.3D. The L + M channel will contribute two lines with negative slopes, and the L - M channel will provide two lines with positive slopes. In that case we should see a diagonally oriented, rectangular contour, with its dimensions determined by the sensitivities of the L + M and L - M channels. Figure 12.3D shows two such rectangles, one expected with stimulus conditions in which the L + M channel is more sensitive than the L - M channel, and the other for conditions for which the L - M channel is more sensitive than the L + M channel.

To summarize, extending what we said earlier, the sizes and orientations of detection contours in cone contrast space give us information both about the identities of the postreceptoral channels
Figure 12.2: Summation squares in two L, M cone spaces. A. Predicted detection contours in threshold normalized L, M cone contrast space. Predictions are shown for summation, subtraction, and independence. B. Method for measuring modulation thresholds in L, M cone contrast space. Thresholds are measured for modulations along many axes through the origin. Each threshold is indicated twice, by the open and closed circles on each axis. A line connecting the data points forms a detection contour in cone contrast space.
Figure 12.3: Predicted detection contours in L, M cone contrast space. Predictions are shown for summation (A), subtraction (B) and independence (C). The less sensitive the channel, the farther apart the predicted contours, as shown by the dotted lines. D. Predictions for a system composed of the combination of two channels, L + M and L - M. The L + M channel contributes the regions of negative slope, and the L - M channel contributes the regions of positive slope. As the sensitivity of the two channels vary, the contour changes size, and changes orientation from the positive to the negative diagonal.
and about their relative sensitivities. For a system made up of two independent channels L and M, we should always see a square or rectangle that maintains its orientation parallel to the L and M axes. The less sensitive the L and M channels the larger the observed square or rectangle should be. But for a system made up from the joint presence of L + M and L - M channels, the observed contour should be oriented diagonally with respect to the L and M axes. The more sensitive the L + M channel and the less sensitive the L - M channel, the more the detection contours will be elongated parallel to the negative diagonal. In contrast, the less sensitive the L + M channel and the more sensitive the L - M channel, the more they will be elongated parallel to the positive diagonal.

One final but important step. Under the theoretical approach we have been using, the exact orientations of the detection contours will be determined by the cone input weights to the L + M and L - M postreceptoral channels. [Work out some numerical examples.] But notice that if the cone input weights are constant and only the sensitivities of the two channels change with stimulus parameters, the axes of empirical detection contours will only lengthen and shorten, as shown in Figure 12.3D, but never change their orientations. Or reversing the logic, sets of diagonally oriented detection contours that stretch and shrink, but never change their orientation, provide evidence for a set of postreceptoral channels with fixed cone input weights – a consistent postreceptoral code. Sets of detection contours that rotate reveal that the cone input weights vary with stimulus parameters, and no single weighting scheme prevails across stimulus conditions.

12.3.3 Curve fitting: Ellipses and superellipses

In mathematical models, detection contours like those predicted in Figure 12.3D are often fitted with ellipses, like that shown in Figure 12.4A. The formula for an ellipse that is aligned with the axes is:

\[
\left( \frac{x}{a_x} \right)^2 + \left( \frac{y}{a_y} \right)^2
\]

where the variables \(a_x\) and \(a_y\) represent the lengths of the two axes of the ellipse. An ellipse corresponds to the Euclidian summation rule discussed with the summation square. The circle in the summation square analysis becomes an ellipse when the two axes are not scaled relative to a threshold.

If the data are too square at the corners to be well fit by an ellipse, one can use a generalization of an ellipse known as a superellipses (Figure 12.4B). For a superellipse, the absolute values of the two terms are taken, and both terms are raised to the power \(\beta\) before being combined:

\[
\left| \frac{x}{a_x} \right|^\beta + \left| \frac{y}{a_y} \right|^\beta.
\]

Such superellipses correspond to the case of probability summation previously presented in the context of the summation square.

The equations given thus far are for cases where the ellipses align with the cone axes. As already introduced, next generalization is to allow for mechanisms that linearly recombine the cones such as the Boynton code. This generalization predicts ellipses of various orientations as shown in Figure 12.1. In Figure 12.1, Panel A shows an ellipse oriented along the L+M and L-M axes rather than
Figure 12.4: Detection contours predicted from (A) an ellipse and (B) a superellipse with an exponent of 4. C, D. Sets of detection contours predicted for stimuli with different spatial and temporal frequencies, given fixed cone input weights for the L + M and L - M channels. If the cone input weights are fixed, the detection contours can elongate only along two fixed axes; they cannot rotate.
the cone axes. Panel B shows the more squared off contours are pieces of superellipses ($2 < \beta < \infty$). Similar equations can be written for ellipsoids or superellipsoids in the three-dimensional case.

When ellipses and superellipses are used to fit empirical detection contours in a cone space, the parameters have ready theoretical interpretations. The major and minor axes of the best-fitting ellipse are identified with the isolation axes of two channels of the postreceptoral color code. The major axis will coincide with the isolation axis along which the contrast threshold is highest – the less sensitive channel – under the conditions tested. The minor axis will coincide with the isolation axis of the more sensitive channel. The relative lengths of the axes correspond to the relative sensitivities of the two channels.

This interpretation requires a grain of salt, however. It is not very interesting if each stimulus condition yields a different set of channels. It is more interesting when a set of conditions yields the same set of postreceptoral channels. Referring back to Panels C and D in Figure 12.3, detection contours aligned with the axes of a particular space reveal independent detection by the channels defined on those axes, whereas diagonally oriented contours reveal combinations of channels that perform summation and subtraction of the signals represented on the axes.

Thank you for being interested enough to follow through this long section on theory! Sometimes it happens that the theoretical treatment takes a long time to work out, but it’s worth it because when the alternate predictions are clear enough, the data discriminate very directly among theories.

### 12.3.4 Empirical detection contours

We are now ready to look at some empirical detection contours. In 1996, Marcel Sankeralli and Kathy Mullen presented an extensive study of detection contours, which they plotted in two-and three-dimensional cone contrast spaces. They varied the spatial and temporal parameters of the stimuli, in order to see whether or not the orientations remained consistent over these variations. Three combinations of spatial and temporal frequencies were used: 1 cy/deg, 0 Hz; 0.125 cy/deg, 0 Hz, and 1 cy/deg, 24 Hz. (For 0 Hz, the stimulus was ramped slowly on and then off in time.) We will use their data to illustrate the shapes and orientations of real discrimination contours, and interpret them in terms of the models outlined above.

### 12.3.5 Plots in 2-D color spaces

Detection contours plotted in an L, M cone contrast space are shown in Figure 12.5. For the data in Figure 12.5A, the spatial frequency was 1 cy/deg, and the temporal frequency was 0 Hz. The result is an elongated, ellipse-like contour, oriented approximately along the positive diagonal. These data provide evidence that for these stimuli, modulations in L, M cone contrast space are detected by a combination of two postreceptoral channels – a relatively more sensitive L - M channel and a relatively less sensitive L + M channel. The orientation of the detection contour close to the positive diagonal means that the L - M channel determines the detection thresholds along the long sides of the contour. This result is consistent with the Boynton code.

For the data in Figure 12.5B, the spatial frequency was reduced to 0.125 cy/deg, and the temporal frequency was kept at 0 Hz. The result is another, larger ellipse-like contour, still oriented

---

8Since detection contours are often well fit with ellipses, these contours could be called *detection ellipses*. They are also sometimes called *discrimination ellipses*, although this term is appropriate only when discriminating two lights rather than detecting a single light on a background.
Figure 12.5: Empirical detection contours in L, M cone contrast space. The three contours were measured with stimuli of different combinations of spatial and temporal frequency. A. 1 cy/deg, 0 Hz. B. 0.125 cy/deg, 0 Hz. C. 1 cy/deg, 24 Hz. A and B consistently suggest a more sensitive L - M channel, and a less sensitive L + M channel. C suggests that with the shift to a high temporal frequency stimulus, the L + M channel has the higher sensitivity. The sizes of these three panels vary to keep the size of the contrast change roughly constant (0.01 vs. 0.05, vs. 0.1). [Modified from Sankeralli and Mullen (1996). A: from Fig. 4, p. 909; B: Fig. 6, p. 911; C: Fig. 8, p. 912.]
12.4. PARADIGM 2: ADAPTATION-NEAR-THRESHOLD

approximately along the negative diagonal. Both channels are less sensitive to the lower spatial frequency, but the L-M channel is still the more sensitive of the two. The unchanged orientation of the contour reveals unchanged weights of the L and M cones in the L-M channel across the change in spatial and temporal parameters.

For the data in Figure 12.5C, the stimulus was again 1 cy/deg, but the temporal frequency was increased to 24 Hz. As a result, the detection contour gets larger, suggesting that the sensitivities of both channels are greatly reduced under these stimulus conditions. The contour also flips to an orientation along the negative diagonal, suggesting that now the L+M channel is the more sensitive. It is such cases with higher temporal (or spatial) frequencies that the luminance channels get their reputation as the more sensitive channel (see section on photometry below).

To look for an S-(L+M) channel, Sankeralli and Mullen shifted to another plane of color space. Since they were looking for the interaction of an S cone signal with a summed L+M cone signal, they looked in a plane whose axes are defined by S and L+M. Alignment of the detection contour with these axes would suggest the presence of both kinds of signals, but argue against their interaction, whereas alignment of the contour along the positive diagonal would suggest the subtractive interaction of these two signals. A detection contour for a spatial frequency of 0.125 cy/deg and a temporal frequency of 0 Hz, plotted in an S vs. L+M space, is shown in Figure 12.6A. The orientation of the contour along the positive diagonal reveals an S-(L+M) channel as stipulated in the Boynton code.

And just to show you that not all interactions happen; suppose you predicted the presence of an S-(L-M) channel. You would look in an S vs. L-M plane, and predict that the empirical detection contour would be oriented along the major diagonal. Data for a spatial frequency of 0.125 cy/deg and a temporal frequency of 0 Hz are displayed in this plane of color space in Figure 12.6B. The detection contours are oriented virtually parallel to the axes – these two channels do not interact and no S-(L-M) channel is seen.

In summary, Sankeralli and Mullen’s data show how the summation-near-threshold paradigm can be used to constrain models of post-receptoral color codes. Most of the available data on detection contours, from their lab and others’, are in good general agreement with the description provided by the Boynton model. This is one of the major reasons that many visual scientists now take the Boynton code as a useful model of the early postreceptoral color code.

12.4 Paradigm 2: Adaptation-near-threshold

12.4.1 Adaptation and cross-adaptation

A second paradigm used to reveal the characteristics of postreceptoral coding is that of adaptation-near-threshold (Graham, 1989). The adaptation-near-threshold paradigm has been widely used in many areas of visual science. To introduce it, we will use the hypothetical example of stimuli composed of 540 and 650 nm lights.

Our schematic adaptation-near-threshold experiment is shown in Figure 12.7. The experiment consists of two steps. First, as shown in Figure 12.7A, we set up a pre-adaptation condition, and measure detection thresholds for two different stimuli – here, test spots of 540 and 650 nm light against a dim “white” background field. Second, as shown in Figure 12.7B, we superimpose each of two adapting fields – here, fields of either 540 or 650 nm – on the “white” background field, and re-measure the thresholds for the 540 and 650 nm test spots. This gives us two same-wavelength
Figure 12.6: Empirical detection contours involving S cone inputs. A. Data collected in an S vs. L + M plane, designed to detect an S - (L + M) channel. The data fall along the positive diagonal, indicating an S - (L + M) channel. B. Data collected in an S vs. L - M plane, designed to detect an S - (L - M) channel. The data align with the axes of the plot, revealing two independent detection channels, and no interaction between them. Thus, no S - (L - M) channel is seen. (The square root of two on the abscissae is a scaling factor that can be ignored for our purposes.) [Modified from Sankeralli and Mullen, 1996, Fig. 6, p. 911.]
12.5. MEASURING FLICKER ADAPTATION IN COLOR SPACE

Various possible outcomes are schematized in Figure 12.7C. The left and right panels show adaptation to 540 and 650 nm respectively. The thresholds measured against the white background field, normalized to a value of 1, are shown by the dark symbols. Now, since the chromatic adapting fields increase the total background intensity, and we know that higher intensity backgrounds elevate detection thresholds (Chapter 10), we would expect threshold elevations when the chromatic adapting fields are introduced. Thus, we would routinely predict threshold elevations for the same-wavelength conditions.

A wider range of predictions are appealing, however, for the cross-adaptation conditions. Three different predictions are shown by the set of three open symbols in each of the cross-adaptation conditions in Figure 12.7C. At one extreme, we might believe that all of the stimulus fields are detected by a single visual channel (for example, a single photoreceptor type). In that case, wavelength can’t matter, and we predict complete cross-adaptation, as shown by the top symbols. But at the other extreme – and here’s the point – we might believe that the 540 and 650 lights are processed by two entirely separate and independent channels within the visual system. If so, then an adapting field processed by one channel would have no effect on the detection threshold for test stimuli processed by the other channel. In that case we predict that no cross-adaptation will occur, as shown by the bottom symbols. And if stimuli of the two wavelengths are processed by two different channels that are only partially independent, partial cross-adaptation might also be predicted, as shown by the middle symbols.

To summarize, in an adaptation-near-threshold experiment, two different stimuli (or stimulus dimensions) are used. The pattern of cross-adaptation between the two kinds of stimuli yields information about the independence or interdependence of processing of the two stimuli. In particular, cross-adaptation is consistent with the conclusion that the two stimuli are processed by the same or only partially separable visual channels. But more interestingly, a failure of cross-adaptation is consistent with the conclusion that the two stimuli are processed by two separate processing channels – physiologically, separate kinds of neurons – within the visual system.

12.5 Measuring flicker adaptation in color space

12.5.1 Light and contrast adaptation

In Chapter 10 we discussed light and dark adaptation. The term adaptation was used to describe the large changes in detection thresholds that occur in the presence of background fields of various luminances (light adaptation), and for many minutes after the termination of a high intensity background field (dark adaptation); and also to refer to the physical and physiological mechanisms that cause these changes in thresholds. Here, we again use the term adaptation to indicate an elevation of a detection threshold. But rather than manipulate the prior state of the mean light level, we manipulate the prior state of contrast. In general, this is called contrast adaptation.

The mechanisms for the two kinds of adaptation cannot be the same, because experiments with contrast adaptation maintain a constant mean luminance. That is, it provides no change in time-average luminance, and cannot “light” adapt the visual system in the old sense of the term. Rather, some mechanism that responds to modulation per se must have been reduced in sensitivity by being exposed to the flicker of the adapting field, and must recover only gradually over time.
Figure 12.7: The adaptation-near-threshold paradigm: Adaptation and cross-adaptation. A. In the first phase of the experiment, thresholds are measured for 540 and 650 nm test spots against a dim 'white' background field. B. In the second phase, 540 and 650 nm chromatic adapting fields are introduced, and thresholds are measured for both test spots against both chromatic adapting fields. C. Predictions. Thresholds should be routinely elevated in the same-wavelength conditions. But three different outcomes are possible for the cross-adaptation conditions. Complete cross-adaptation suggests that both wavelengths are processed by a single processing channel; partial cross-adaptation suggests two partially independent channels; and an absence of cross-adaptation suggests two independent channels.
12.5. MEASURING FLICKER ADAPTATION IN COLOR SPACE

More particularly, we need to introduce the phenomenon of flicker adaptation. Again for concreteness, suppose that a subject views a homogeneous “white” adapting field, upon which a small “white” test spot – a brief luminance increase – can be superimposed. In the first phase of the experiment, the subject sets a detection threshold for the test spot. In the second phase, the adapting field is flickered (modulated in time) at a high contrast, above and below the mean luminance used in phase 1 of the experiment. The subject is asked to stare for a minute or two at the flickering adapting field. And finally, in the third phase, the subject is asked to reset the detection threshold for the test field. The question is, will the threshold for the test field be changed by exposure to the high contrast flicker of the adapting field?

Typically, the exposure “adaptation” to the modulating adapting field will elevate the detection threshold for the test stimulus, sometimes by as much as 0.5 log units (a factor of 3). The subject’s sensitivity recovers only gradually, perhaps over the ensuing 20 to 30 seconds. A common way to think about this result is to assume that somewhere within the visual system, neurons that respond to the flickering field become fatigued (or are reduced in sensitivity, or have their dynamic ranges shifted upwards) during the adaptation interval, so that their detection thresholds are elevated for some time after exposure to the flickering field.

Now the next interesting question is, is there cross-adaptation? That is, does flicker adaptation generalize across stimulus dimensions? For example, if the adapting flicker is a luminance modulation, would it elevate the thresholds only for luminance-modulated stimuli, or would it also elevate the thresholds for stimuli defined solely by chromatic changes, such as a change of the test field from “white” to an isoluminant “red” and back again? If the luminance-modulated adapting field elevates the threshold for the white-to-red-to-white test pulse, and vice versa, one can argue that these two stimuli must affect the same neural channel (otherwise they could not interact). Conversely, if there is no cross-adaptation, one can argue that luminance-modulated and chromatically modulated stimuli are detected and processed by two separate, independent channels within the visual system.

12.5.2 Color codes: Evidence from differential cross-adaptation

A classic flicker adaptation experiment in color vision was carried out by John Krauskopf, David Williams, and David Heeley in 1982. Their experiment is schematized in Figure 12.8. Krauskopf and his colleagues reasoned that if the adapting light is modulated along one isolation axis in post-receptoral channels space, detection thresholds on that axis should be elevated; but by hypothesis, that adapting light creates no variation at all in the signal in the other two channels. Thus, if the choice of the model is right, these channels cannot become adapted, and their contrast thresholds cannot be changed. And conversely, adaptation along a second or third isolation axis should have no effect upon the threshold along either of the other two isolation axes. In other words, there should be no cross adaptation between any two true isolation axes.

On the other hand, suppose the modulation of the adapting light occurs along an axis in between two isolation axes. The adapting light would create modulated signals in both of the two channels. In consequence, one should see cross-adaptation: since both channels should adapt, the subsequent thresholds for the test pulse should be elevated on both axes, and all axes in between. And an adapting light that modulates the signals in all three channels should elevate thresholds for pulses on all axes of color space.

In sum, this experiment adds another layer to the logic of the adaptation-near-threshold paradigm.
Figure 12.8: Paradigm for Krauskopf et al.'s experiment. A. First, a plane is defined by choosing two axes, A and B. Initial threshold measurements are made for pulses away from the origin along many axes in this plane. These thresholds are normalized to trace out a circular detection contour, as shown. B. Next, the subject is adapted to high contrast modulation (flicker) along any one of four axes in the chosen plane, as shown. Finally, thresholds are remeasured along all of the original axes. C. Possible outcomes. Assume that thresholds are measured after adaptation to flicker along Axis A, as shown by the arrows above and below the figures. The data could show cross-adaptation – threshold elevations on all axes (right panel); or no cross-adaptation – threshold elevations across only a narrow range of axes, specifically excluding Axis B (left panel).
12.5. MEASURING FLICKER ADAPTATION IN COLOR SPACE

It is designed around the premise that if there is a three-channel code, there will be a differential cross-adaptation among different sets of axes. That is, there should be a failure of cross-adaptation among a unique set of three and only three axes in color space, and cross-adaptation among all other sets of axes. Turning the argument around, a particular pattern of results – the absence of cross-adaptation among a unique set of three axes, and the presence of cross-adaptation among all other sets of axes – suggests that the three axes that do not cross-adapt are the isolation axes of the postreceptoral color code. Importantly, then, this experiment provides our second potential line of psychophysical evidence concerning the post-receptoral color code.

How did it come out? The results of Krauskopf et al’s experiment were that no cross-adaptation occurred among the L + M, L - M, and S - (L + M) axes, whereas cross-adaptation did occur among sets of intermediate axes. Sample data are shown in Figure 12.9 for the case of cross-adaptation between axes within the isoluminant plane. The results thus suggest that three channels, L + M, L - M, and S - (L + M) exist as a postreceptorial color code, and that this code is in force at the physiological locus at which flicker adaptation has its effect. This experiment, published in 1982, in fact provided one of the major early sources of support for the Boynton model.

To further appreciate these results, consider another possible outcome. What if the effect of flicker adaptation was specific to a redundant narrow range of colors around the adapting direction? Such a theory is often called a higher-order color mechanism. In that case, one would expect no cross adaptation for any conditions. This was not found for flicker adaptation thus no need to propose such higher-order mechanisms for this case. But it has been found for other situations including color appearance judgments following flicker adaptation (Webster and Mollon, 1991).

---

Krauskopf et al called the independent axes discovered in their experiment the cardinal axes of color space. Another term that DT uses is privileged axes. The idea is that the choice of axes is to some extent unique. For this example, that refers to obtaining the pattern of cross adaptation with the chosen axes but not with an alternative choice.
12.6 Summary of evidence for the Boynton code

In this chapter, we have examined two psychophysical paradigms – summation-near-threshold, and adaptation-near-threshold – designed to reveal the postreceptoral color code. Both paradigms yield data consistent with the Boynton code. Two very different psychophysical paradigms thus reinforce each other. We should emphasize at this point, however, that much untidiness remains. The results from different laboratories on different individual subjects, with the use of different experimental parameters and stimulus paradigms, are more variable than ideal. In general, the paradigms that reveal specific coding schemes, reveal the Boynton code; but other paradigms reveal less unique sets of privileged axes. The data can be reconciled post hoc by assuming that some of the experiments are controlled by a level of the system at which the Boynton code is in force, whereas others are controlled at other levels. For example, as cited above, the detection thresholds following flicker adaptation followed the predictions of the Boynton code but similar appearance experiments did not.

In DT’s view, this situation is not as bad as it might seem – it’s just that the experiments we are discussing are at the current forefront of science. Remember that trichromacy was known for a century or so, during which many creative but incompatible estimates of the spectral sensitivities of the three cone types were proposed. Eventually the data from different paradigms converged closely on a single set of cone spectra (Figure 7.2). History repeats itself. We are now in the messy process of coming to consensus on the cone inputs to the three channels of the early postreceptoral code. The visual system presumably contains at least a few later levels with additional color code transformations, and different levels will control the data when different paradigms and stimulus parameters are used. The task ahead is to make a consistent multistage model that incorporates all of the data.

12.7 Photometry revisited

12.7.1 Degenerate detection contours: Tubes and pancakes

Finally, let’s return to the question of detection contours, and reconsider them in postreceptoral channels space. These ideas will lead us to propose an explanation for the facts about photometry, introduced in Chapter 3.

Now let’s look at some theoretical examples of degenerate cases of detection contours in three-dimensional postreceptoral channels space, shown in Figure 12.10. Suppose we were to try to measure a detection contour, but we used stimuli to which one of the channels – say the S - (L + M) axis – is very insensitive, as shown in Figure 12.10A. The detection contour would elongate along the axis of insensitivity. In the extreme, suppose the subject simply could not detect the stimulus modulation along that axis, even at the highest modulation available with the equipment being used to make the measurements. In that case, there would be no way to measure the contrast threshold on that axis – all we could do would be to indicate that at the highest available contrast, the stimulus remained undetectable. We will indicate such an option with outward pointing arrowheads, as seen on the S - (L + M) axis in Figure 12.9A. Assuming the contrast thresholds on the other two axes can be measured, the detection contour would take the form of an open ended tube, oriented along the S - (L + M) axis. And conversely, an empirical detection contour that degenerates into a tube suggests that the visual system has only two rather than three functional channels available under...
the conditions tested.

Now suppose the visual system is even more limited. Suppose that we use stimuli that (say) only the $L + M$ channel can detect. In that case, there would be no measurable threshold along the $L - M$ axis, nor along the $S - (L + M)$ axis, nor along any other axis in the isoluminant plane. In fact, the stimulus would be detected only when it made a detectable signal in the $L + M$ channel. In that case, as shown in Figure 12.10B, the detection contour would degenerate to a pancake – two planes parallel to the isoluminant plane, one above and one below it, separated by a distance equal to the contrast threshold on the $L + M$ axis. And conversely, a detection contour that degenerates into a pancake suggests that only a single functional channel is present under the conditions tested. Moreover, the detection contours bracket the null plane of that channel.

An example of some data that approximate pancakes is provided by an experiment in which DT served as a subject. In this experiment, Teller and Lindsey (1993) used moving sinusoidal grating stimuli, modulated through “white” along many different axes within and tilted out of the $V(\lambda)$-defined isoluminant plane. On each trial, the grating moved either upward or downward, and the subject’s task was to report its direction of motion in a forced-choice task. The experiment was repeated along four different chromatic axes.

The results for subject DT are shown in Figure 12.11. The motion threshold could not be measured along a small range of axes with tilts very close to the isoluminant plane. It became measurable for slightly greater tilts of the axis, and then decreased rapidly as the tilt approached the achromatic axis. For each of the four different chromatic axes tested, the data traced out two parallel lines, making a pancake overall. Lindsey and Teller called the space within the pancake the motion dead zone – the region of three-dimensional color space within which no motion was seen. Their data suggest that for the stimuli and task used, there is only one functional channel – the $L + M$ channel – and that DT’s isoluminant plane differs slightly from the standard isoluminant plane.

### 12.7.2 Photometry and the isoluminant plane

In Chapter 3, we discussed the topic of photometry. We noted that three different methods – flicker photometry, minimally distinct borders, and motion photometry – all yield the same or highly similar photopic spectral sensitivity curves, and we wondered why this might be so. At this point we are ready to unite our earlier discussion of methods of photometry with the present discussion of three-dimensional postreceptoral channels spaces, and provide a simple model of why photometry works. Take flicker photometry as an example. Let us assume (as is true) that the two chromatic channels are limited in temporal resolution, and cannot follow 15 Hz flicker. In consequence 15 Hz flicker is detected solely by the luminance channel. Let’s carry out a thought experiment on the detection of 15 Hz flicker, along all of the different axes through the white point at the origin of DKL space.

What we expect to find is a pancake, with its crusts parallel to the isoluminant plane. There are two factors at work here. First, neither of the two chromatic channels can detect the flicker, so the $L + M$ channel is the only available possibility. Second, when we find the isoluminant plane – the null plane for the $L + M$ channel – and flicker two lights from within it against each other, the $L + M$ channel no longer detects the flicker either. By hypothesis it was the only channel that could follow fast flicker, so when that channel is nulled, no flicker can be seen. Or reversing the logic, when the subject reports that he perceives no flicker, we have put the stimulus modulation
Figure 12.10: Degenerate cases: Tubes and pancakes. A. If one channel (say the S - (L + M) channel is so insensitive that no detection threshold can be measured, the detection contour degenerates into a tube along that axis. B. If two channels (say the two chromatic channels) are so insensitive that their thresholds cannot be measured, the contour degenerates into two parallel planes (a pancake). The orientations of the degenerate contours reveal the postreceptoral color code.
12.7. PHOTOMETRY REVISITED

Figure 12.11: A pancake measured in a direction-of-motion task. Stimuli were modulated along axes through the origin. In the top left panel, the matching letters (a, a; b, b; etc.) show the two representations of the same contrast threshold. The subject’s task was to judge the direction of motion of the stimulus. The four panels show four cuts through the L + M axis, along the L + M axis (top left), the S - (L + M) axis (top right), and two intermediate axes (bottom panels). In each case the contrast threshold became too large to measure at a small range of axes of modulation near the isoluminant plane, as indicated by the arrows. Thus, all four data sets together trace out a pancake that nearly coincides with the isoluminant plane. This subject’s isoluminant plane is tilted slightly from the isoluminant plane predicted from \( V(\lambda) \). [Modified from Teller and Lindsey (1993, Fig. 3, p. 1328).]
within the null plane for his L + M channel.

A similar argument can be made in the spatial domain in the case of the minimally distinct border technique. Let us also assume (as is true) that the chromatic channels cannot signal the presence of high spatial frequencies. The L + M channel, as we know, can signal spatial frequencies up to 60 cycles/degree. Now, make the reasonable assumption that the perception of a sharp border depends on having activity in a channel that responds to high spatial frequencies. Then sharp borders can be perceived whenever there is a signal in the L + M channel – that is, when we modulate out of the isoluminant plane. But when the only channel that can do the analysis – the L + M channel – is nulled by modulating within the isoluminant plane, the perceptual sharpness of the border is lost.

A similar but more imaginative story can also be told about the motion minimization technique. The analysis of motion is a complex computational task, which we will discuss further in Chapter 20. But for the moment, let’s adopt the assumption that the analysis of motion is done solely on signals in the L + M channel. In that case, the perception of motion, like the perception of fast flicker or sharp borders, would be lost for stimuli confined to the isoluminant plane.

From these examples we can extract a general theoretical argument. The argument is that there are psychophysical tasks – for example, following fast flicker, or processing the fine details of a border, or analyzing the direction of motion – that for whatever reason are accomplished by calculations on only the signal in the L + M channel, and cannot be sustained by signals in the chromatic channels. The subject’s capacity to do these tasks will therefore fail when the only channel that can do the task – the L + M channel – is disabled, by using stimuli that fall in its null plane.

We can now see that in the theoretical approach we have been developing, the science of photometry has to do with finding tasks that only the L + M channel can do, and then nulling that channel. The photometric system thus defines stimuli of different wavelength composition that fall in the null plane for the L + M channel. Thus, in principle, any task that the L + M channel is good at and the chromatic channels are bad at, could potentially be used as a basis of photometry. Flicker, border distinctness, and motion are the three most commonly used photometric tasks, but there are others as well.

In sum, at the beginning of this chapter we stated that since the whole human subject can monitor all of her channels at once, and detect a stimulus with any channel that responds, null planes in three dimensional color spaces do not usually reveal themselves for the subject as a whole. But we now see that photometry points out the exception to this statement. If we design or encounter stimulus conditions in which only one of the available channels can respond to the stimulus and do the analysis required by the task, and then adjust the axis of stimulus modulation to fall within the null plane of that channel, our subjects should fail to detect that stimulus, or to analyze its properties. Such cases are probably rare outside the laboratory, but in the laboratory our diabolical cleverness can be put to use, and our stimuli can be custom designed for the purpose. The practice of photometry illustrates a set of cases in which, in the context of this theory at least, we do determine a null plane with psychophysical procedures.

12.8 Summary

In this chapter we have pursued two goals: to introduce two general psychophysical paradigms used to explore the properties of postreceptoral coding, and to work these two paradigms through for
the case of the postreceptorial color code. The first paradigm, summation-near-threshold, reveals summations and subtractions of cone inputs that conform rather closely to the Boynton code. The second paradigm, adaptation-near-threshold, reveals a unique independence of processing between stimuli that isolate the three channels of the Boynton code, and it therefore similarly provides evidence that the early postreceptorial code conforms quite closely to the Boynton code. In the next two chapters we consider support for this color code for the ganglion and LGN physiology.

Finally, we returned to the topic of photometry introduced in Chapter 3, and placed it within the context of detection contours in three-dimensional color spaces. In this context photometric tasks are seen as tasks that the visual system, for whatever reason, can do only with signals in the L + M channel, and not with signals in the L - M or S - (L + M) channels. In the cases of flicker photometry and minimally distinct borders, these limitations come about because the high temporal and spatial frequencies involved exceed the spatial and temporal resolution of the chromatic channels. In the case of motion photometry, we introduced a novel argument – that the analysis of motion might be carried out on only signals in the L + M channel. In all three cases, our new interpretation of why photometry works is that photometric techniques correspond to circumstances in which only the L + M channel is sensitive enough to detect the stimulus, and photometric matches reveal the null plane of the L + M channel.