Vision Res. Vol. 22, pp. 635 to 643, 1982
Printed in Great Britain

0042-6989/82/060635-09$03.00/0
Pergamon Press Ltd

WHAT CAUSES TRICHROMACY? A THEORETICAL ANALYSIS USING COMB-FILTERED SPECTRA

H. B. BARLOW
Physiological Laboratory, University of Cambridge, Downing Street, Cambridge CB2 3EG, England

Abstract—For colour vision, the task of the eye is to discriminate different distributions of energy over the spectrum. This is usually treated as a problem in the wavelength domain, analogous to treating spatial resolution in terms of spatial positions in the image. What is attempted here is a treatment of colour vision in terms of the system's responses to spectral energy distributions that are sinusoidal functions of wavelength. These are called comb-filtered spectra, and the treatment is analogous to that of spatial vision in terms of spatial sinusoids. This gives some insight into the reasons for trichromacy, the advantages of oil droplets, and the narrow separation of the red and green mechanisms. It is also shown that the absorption spectra of photosensitive pigments are superimposable if plotted as a function of the fourth root of wavelength.

William Rushton was a marvellous person to discuss ideas with. I discovered this when, as a research student, I demonstrated in the "Special Senses" practical classes he ran for second year physiologists. He had not been responsible for designing the class, much of which was "borrowed" from Psychology and had originally been set up by G. C. Grindley and A. F. Rawdon-Smith, perhaps with an earlier influence from Hamilton Hartridge. Neither of us knew much optics, and I recall our prolonged attempts to explain retinoscopy to each other, and occasionally to a medical student. The task of explanation was complicated by the variety of equipment available. There were one or two standard concave mirrors of the usual type, but most of them were plane mirrors, so as well as explaining why the reflections in the pupil moved the way they did, as often as not we also had to explain why William and I were directly contradicting each other.

We came across a good many mysteries of that sort from which I certainly learned a lot of optics. One of these was connected with the plane retinoscopy mirrors mentioned above. With typical Cambridge extravagance these consisted of Woolworth handbag mirrors from which the backing had been scraped away over a small circle. If you reflect the light from a point source on to a white surface with such a rectangular mirror, it of course has a dark spot in the middle of the quadrilateral patch of light. But if you place your eye behind the hole in the mirror and look at this patch through it, with your eye correctly placed the spot in the middle is seen as bright instead of dark. If I recall correctly, I pointed out this phenomenon to William in the middle of a session, and at first we were completely baffled. William was relentless when faced with such a problem, and by then the medical students knew better than to try to ask a question when his curiosity had been aroused and he was on the chase. After drawing a good many diagrams on the blackboard we had the answer, which is shown in Fig. 1: the point source of light forms an image on one's own retina, and the light leaving the eye through the lens from this image is reflected back into the pupil from the back of the retinoscopy mirror, perhaps aided by incompletely removed silvering. This forms a parasitic image that lies exactly where the dark patch would be. In this way one can inspect an image that results from three transits of the eye's dioptric mechanism and I suppose one might be able to use the method to assess the quality of the retinal image, and perhaps to make the bleaching of retinal pigments directly visible in one's own retina!

I remember another hilarious episode in which we discovered a medical student with a mysterious cyclophoria. When Maddox rods were placed in front of one eye this student stubbornly maintained that the bright line he saw passed exactly through the point of light, but was oblique, not vertical. We spent a good many minutes watching his eye to see if it rotated as the rods were placed in front of his eye, and then we both became suspicious: the student seemed to be enjoying himself altogether too much and we thought he must be pulling our legs, so we tested him by tilting the rods. "Which way does it tilt?" we asked him, and he proceeded to give appropriate answers for half a dozen or so trials. Defeated, we put the rods down for a moment, then, thinking he could perhaps have seen which way we had rotated the rods on each trial, we started a new series. "Which way?" said William, holding the rods vertical. "To the left" said the student. "What!" said William, "you always said it was to the right before." "Yes", answered the student, "but you've got the plate back to front this time." At this we looked more carefully at the so-called "Maddox rods". This was another example of Cambridge extravagance, namely a 4 × 2 in. rectangle cut from a plate of ridged glass such as is used for glazing lava-

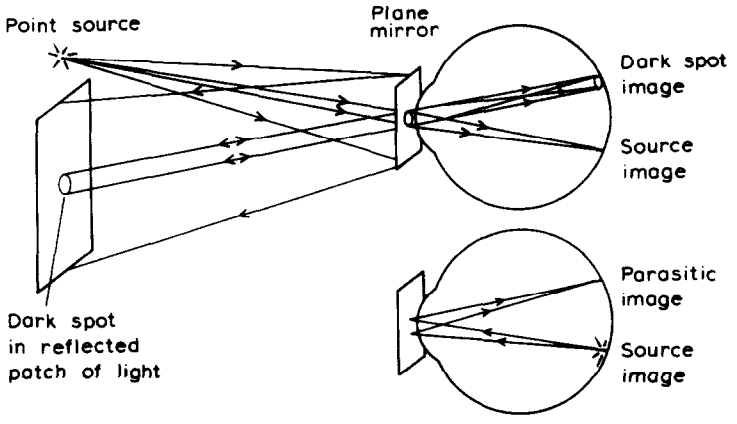

Fig. 1. The patch of light from a mirror with a hole scratched in the backing has a dark spot at its centre, but if one looks at this patch through the hole in the mirror, with one's eye correctly positioned one sees a bright spot at the centre instead. This is caused by a parasitic image of the light source formed by triple passage through the eye's optics. The top section shows the formation of the source image and of the expected image of the dark spot. Below is shown the formation of the parasitic image, which lies at the position of the image of the dark spot.

tory windows and the like, but this particular one had been cut with the ridges some 15° oblique to the edge of the rectangle. I always think of this episode when I read papers, often using thorough statistical analysis and a commendably large number of subjects or animals, in which the author gives no evidence at all of having bothered to inspect the test patterns allegedly produced by his equipment.

The following discussion actually arose from comments by John Mollon about a passage I had written about the merits of narrow and broad spectral sensitivity curves for our forthcoming textbook (Barlow and Mollon, 1982). It is very much the kind of idea it would have been enjoyable and fruitful to discuss with William Rushton, though I have to admit that his gift for spotting the defects and limitations of an argument would have been more likely to have inhibited than promoted its publication.

COMB-FILTERED SPECTRA: SINUSOIDAL DISTRIBUTIONS OF SPECTRAL ENERGY*

Hitherto almost all serious analysis of colour vision has aimed at using narrow spectral lines as stimuli. Often, for practical reasons, these lines have not been as narrow as the experimenter would have liked, but the conceptual framework had been in terms of spectral bands, and spectral weighting functions. This type of analysis may be likened to the analysis of spatial vision in terms of localized spots and patches, but since the introduction and popularization of sinusoidally varying stimuli (de Lange, 1952; Schade, 1956; Campbell and Robson, 1963) vision researchers have become accustomed to the merits of analysis in terms of frequency transfer characteristics and contrast sensitivity functions. Is there any potential in the analysis of colour vision by comb-filtered spectra,* i.e. stimuli whose energy varies sinusoidally with wavelength?

Figure 2 shows that such stimuli have an honourable history. It is from Newton's *Optiks*, and illustrates an experiment in which he inserted a comb into a spectrum that was subsequently recombined by the lens into a small patch of light. Without the comb this appeared white, but the comb blocked part of the spectrum and thus made it coloured, the colour changing as the comb was slowly moved to and fro. When it was moved fast the patch appeared white again and thus Newton was able to show that white light results from the successive presentation of the spectral colours, just as it does from their simultaneous presentation. From the dimensions that Newton gives one finds that one cycle of the comb occupied almost the full length of the spectrum, and he does not report any observations with finer combs. In the present paper, I consider what would happen if a comb with many sinusoidally shaped teeth was moved slowly across a spectrum to produce comb-filtered spectra of various comb-frequencies and phases. Their amplitude of modulation might also be varied, and it will be interesting if this gives evidence of colour opponency, as the equivalent spatial experiment gives evidence of lateral inhibition. However,

*The term comb-filtered spectrum is used because such stimuli could be produced by inserting a comb into a spectrum and then recombining it, as shown in Fig. 2. Comb-filtering has also been used to describe sound stimuli in which energy varies periodically with sound frequency. Furthermore, a possible alternative term, "colour gratings", has been pre-empted to describe spatial gratings in which colour, but not luminance, varies with position. The frequency and wavelength of the sinusoidally modulated energy distributions will be referred to as the "comb-frequency" and "comb-wavelength" to avoid confusion with spectral frequencies and wavelengths. These terms will also be used to describe the set of sinusoidal spectral energy distributions into which any arbitrary spectral energy distribution can, by Fourier's theorem, be decomposed.

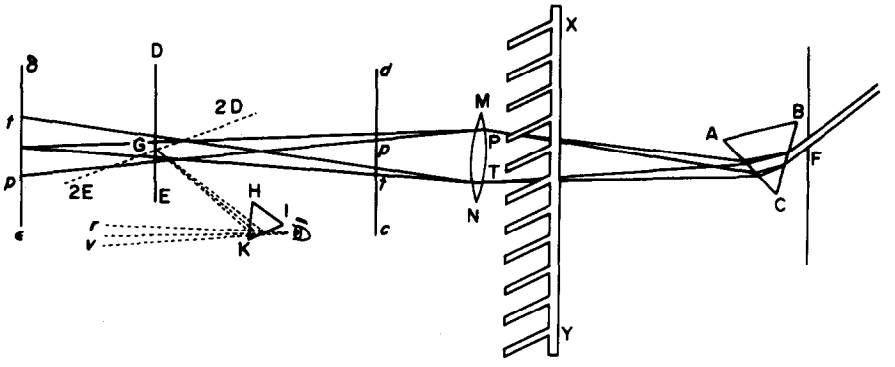

Fig. 2. Newton passed sunlight from a hole in his window-shutter (F) through a prism (ABC) to produce a spectrum on a lens (MN). The lens recombined the spectrum to form a white spot in the plane DE. When a comb (XY) was placed in the spectrum parts of it were blocked and the white spot became coloured, the colour changing as the comb was moved. However, when it was moved rapidly the different colours appeared in rapid succession and were again combined to give a white appearance. The teeth of his comb were separated by almost the complete length of the spectrum, and he does not report any observations with combs of higher frequency. The question considered in this paper is how high a comb-frequency could be "resolved" by the colour system; that is, what is the highest comb-frequency that would give a non-white appearance that varied with comb-position?

without any experiment I think their theoretical consideration does clarify three points.

The first is the dimensionality of colour space—its familiar trichromacy—and even before that there is an unexpected spin-off. The analysis would be much simpler if all the spectral sensitivity curves were the same breadth, and in seeking a function of wavelength that would bring this about I stumbled on a very simple transformation of the wavelength scale that gives all absorption spectra the same breadth, so that they can be superimposed by a simple lateral shift with surprising accuracy. Unfortunately, this does not work nearly so well for the putative spectral sensitivity curves of the colour mechanisms, probably because of the unequal effects of the corrections for ocular absorption. Returning to trichromacy, sampling theory offers an explanation quite different from the conventional one in terms of three different pigments, but this can be rejected by inspection of the calculated modulation transfer characteristics of the spectral pigments. The margin of rejection is not large, however, and the coloured oil droplets of birds and reptiles, which narrow the receptors' spectral sensitivity curves, may be an adaptation making tetrachromacy more worthwhile.

The second result of the analysis is to resolve my argument with John Mollon: broad curves of spectral sensitivity must cause a loss of discrimination between narrow-band stimuli such as those used for measuring $\Delta\lambda$, though this loss will be less for the broad spectral bands likely to be caused by most natural pigments. Finally, the analysis highlights the curious positioning of the peaks of the three-colour mechanisms; the close proximity of the green to the red would be far from optimal for colour discrimination and it is suggested that it enables the two types of cone to be used without distinction for spatial vision.

RESPONSES TO COMB-FILTERED SPECTRA

Figure 3 shows the spectral sensitivity curves for the blue, green and red systems given by Boynton (1979) from the results of Smith and Pokorny (1975), together with two comb-filtered spectra. It is clear

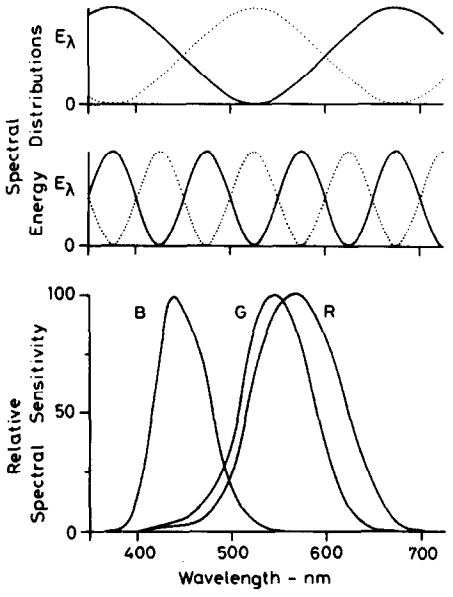

Fig. 3. The top two sections show comb-filtered spectra of frequencies 3.3 cycle/μ and 10 cycle/μ, each at two phases. The lowest section shows the three primaries of Smith and Pokorny (1975) plotted as relative spectral sensitivity curves. It is clear that the colour of the low-frequency comb-filtered spectrum will change with its phase, but less clear whether this will be the case with the high-frequency one. From the Fourier transforms of the spectral sensitivity curves given in Fig. 4 one can see that the red primary will be almost unaffected by the phase of the high comb-frequency stimulus, which will be demodulated to 2.5%, whereas the blue curve will only demodulate it to 25%.

that when the comb-frequency is low the three mechanisms would be excited to different extents than by a uniform, unmodulated, distribution of energy, and they would thus appear coloured. Furthermore, the colour would change as the phase of the grating was changed. In order to make a stimulus of this low comb-frequency look like an unmodulated, uniform, spectral distribution one would obviously have to decrease its contrast considerably.

Consider next the stimulus of high comb-frequency: how can one decide if this would ever appear coloured? The classical colorimetric method would be to work out the integrals $\int V_\lambda E_\lambda d\lambda$ for each of the three V_λs and then see if they differed sufficiently from the values for an unmodulated light to be above threshold. This would then have to be repeated at different phases. However, instead of doing this for all comb-frequencies it is simpler to use the short-cut offered by the Fourier transforms of the spectral sensitivity curves.

Take a single spectral sensitivity curve and imagine the output of its corresponding cone as a function of the phase of a comb-filtered spectrum. This will be a sine wave and what we require is its amplitude as a function of the comb-frequency of the stimulus. By the familiar multiplication theoreom (Woodward, 1953; Bracewell, 1965), the Fourier transform of the output is the product of the Fourier transform of the input times the Fourier transform of the weighting function, so if we consider inputs containing each frequency in turn, the amplitudes of the outputs correspond to those of the Fourier transform of the spectral sensitivity curve. This is exactly similar to regarding an insect ommatidium (Barlow, 1965) or a retinal receptive field (Enroth-Cugell and Robson, 1966) as a filter for spatial frequencies and it is illustrated in Fig. 4.

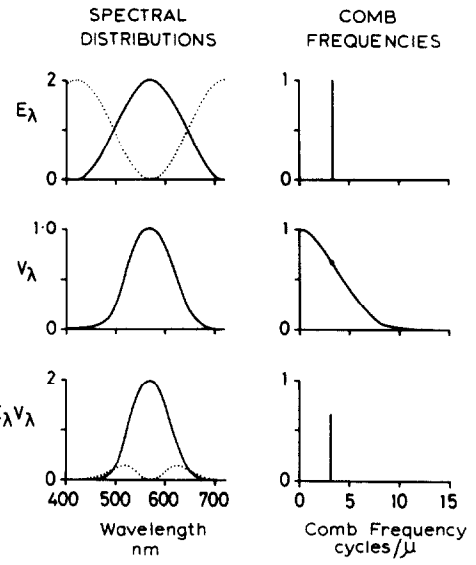

Fig. 4. The steps for calculating the visibility of comb-filtered spectra in the spectral wavelength domain (left) and comb-frequency domain (right). The top pair represent a comb-filtered spectrum in the two domains, two phases being shown at left. The middle row shows the spectral sensitivity curve of the red machanism (Smith and Pokorny, 1975) and its Fourier transform. The bottom row shows the product of the two. The calculation in the wavelength domain must be continued by calculating the integrals of the curves for different phases of the input grating, and calculating the contrast as (Max − Min)/(Max + Min). In the frequency domain the amplitude of the Fourier transform gives the demodulation or contrast loss for all frequencies directly.

Figure 5 shows the Fourier transforms* of the three colour mechanisms [from Boynton (1979) and Smith

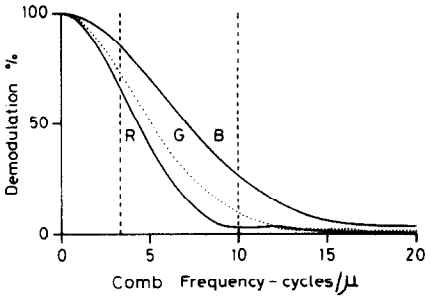

Fig. 5. The Fourier transforms (amplitude) of the three human colour systems from Smith and Pokorny (1975). These show the extent to which the modulation of comb-filtered spectra of different frequencies is decreased by each of the three systems. Note that the low-comb-frequency stimulus of Fig. 3 (3.3 cycle/μ) is demodulated to 66, 73 and 85% by the red, green and blue systems, respectively whereas the higher frequency (10 cycle/μ) goes to 2.5, 9.5 and 26% respectively. Comb-frequencies above 12.5 cycle/μ (i.e. comb-wavelengths below 80 nm) must be very poor at stimulating the colour system simply because the broad spectral sensitivity curves act as low-pass filters demodulating comb-frequencies in this range and higher.

*Fourier transforms and other manipulations of sensitivity curves and absorption spectra were performed on a computer as follows. First the data were entered in a computer file at 10-nm intervals, since the results are usually published or otherwise available in this form. If they did not come down to zero at both ends of the range it was necessary to extrapolate to zero values before Fourier transforms were obtained, but the extent to which this was necessary on the data used in this paper is trivial (of course it is not necessary at all for simple monotonic transformations of the wavelength scale of absorption spectra as in Fig. 6). From the primary file a secondary file was computed using a Lagrange four-point interpolation procedure, these interpolated points being regularly spaced at appropriate intervals on whatever transformed wavelength scale was required (e.g. $\lambda^{1.4}$, λ^{-1}). The first 10 sine and cosine coefficients of the Fourier series for this file were then determined by the appropiate multiplications, using a total range of regularly spaced values of the secondary file chosen to give sufficient points to enable smooth curves to be drawn. For the smooth, gaussian-like, sensitivity curves of the postulated colour mechanisms, 10 coefficients were found to be sufficient. The various stages of the computation were shown on a graphics display, and the adequacy of the transform checked by resynthesizing the sensitivity curves from the coefficients. For determining the demodulation of comb-filtered stimuli it is the amplitude of the Fourier transform that is of interest and these are what are shown.

Table 1. Half bandwidths of red and blue photoreceptor pigments of *Macaca fascicularis* plotted on various scales [from microspectrophotometric results of Bowmaker *et al.* (1980) measured on long wavelength side of peak]

	f(λ)	λ (nm)	1/λ (cm^{-1})	log λ	$\lambda^{1/2}$	$\lambda^{1/4}$	$\lambda^{1/5}$
Red	Peak	567	17640	2.754	23.81	4.880	3.554
	Width	51.7	1475	0.037	1.068	0.1076	0.0626
Blue	Peak	416	24038	2.619	20.34	4.516	3.341
	Width	41.0	2165	0.041	0.981	0.1074	0.0634
Red bandwidth / Blue bandwidth		1.26	0.681	0.902	1.089	1.002	0.987

The top row shows the functions used, the bottom row the ratio of bandwidths on that scale. It will be seen that these two pigments have very nearly the same bandwidth on a scale of $\lambda^{1/4}$. Note that $1/\lambda$ actually increases the change of breadth with position of λ_{max}.

and Pokorny (1975)] of Fig. 2. Note first that they would demodulate the following comb-frequencies by 90%: 13.3 cycle/μm for the blue; 9.9 cycle/μm for the green; 7.7 cycle/μm for the red. As figures for the limits of colour vision one can take the wavelengths at which wavelength discrimination has increased by a factor of five and is rising sharply, namely 435–650 nm, a range of 215 nm (Wright and Pitt, 1935). One can see that there is only just room for 2 cycles of a 10 cycle/μm comb-filtered spectrum, which would be demodulated to 2.5, 10 or 25% by the red, green or blue mechanisms, respectively. Clearly, demodulation occurs even at these low comb-frequencies, but it would be easier to reason about the way the three primaries sample the information in spectral energy distributions if their spectral sensitivity curves were all the same shape, and the following section describes how the absorption spectra of the receptor pigments can be made similar. The reader must, however, be warned that this is a digression from the main argument of the paper, for although the result is interesting and useful in its own right, what is appropriate for absorption spectra turns out to be fruitless as an aid in the analysis of trichromacy.

THE BREADTH OF SPECTRAL ABSORPTION CURVES AS A FUNCTION OF THE FOURTH ROOT OF WAVELENGTH

Dartnall (1953) noticed that the absorption spectra of photosensitive pigments became narrower if λ_{max} was at shorter wavelengths, and he provided a nomogram suggesting they might be the same breadth if a scale of frequency ($1/\lambda$) was used instead. The results of Bowmaker *et al.* (1980) for the red (λ_{max} = 567 nm) and blue (λ_{max} = 416 nm) pigments of cones from *Macaca fascicularis* span a large range of λ_{max}, and I have used these to explore how bandwidth changes according to the function of wavelength used to plot the absorption spectrum. The second row of Table 1 shows the value of the peak of the red pigment, using various functions of wavelength, and the fourth row shows the same values for the blue pigment. Below each peak value is the bandwidth measured to the 50% absorbance point on the long wavelength side (following Bowmaker *et al.*), using the same scale. The sixth row shows the ratio of these bandwidths. It will be seen that red is 26% broader than blue on the linear scale, and 32% narrower on the frequency scale. On the scale of $\lambda^{1/4}$ they are equal, within the accuracy of the measurements, whereas on a scale of $\lambda^{1/5}$ they again diverge substantially.

Figure 6 shows the full spectra of the monkey pigments plotted on the $\lambda^{1/4}$ scale (left), and superimposed (right). It is evident that they fit well, but two remarks are needed. First, the narrowing of the blue pigment on the short-wave side is not to be relied on, for the instrument was being pushed to its limit in this region and the cones may well contain blue-absorbing substances that are not photosensitive pigments. Second, the $\lambda^{1/4}$ rule is unlikely to be exact; the theory of pigment absorption given by Stiles (1948) leads one to expect a constant slope on the $1/\lambda$ scale on the long wavelength side for all pigments independent of their λ_{max}. Lewis (1955) gave a more refined theory that predicted a shallower slope for the green than for the red pigment, and this was shown to be the case experimentally by Brindley (1955). Nonetheless, to a good approximation these monkey pigments have the same shape on the $\lambda^{1/4}$ scale.

Ebrey and Honig (1977) assembled measurements of the absorption spectra of many pigments based on both Vit A_1 and Vit A_2, and showed that bandwidth, expressed in cm^{-1}, decreased as the wavelength of λ_{max} increased; the A_2 pigments showed the same trend but were broader at the same λ_{max}. If an absorp-

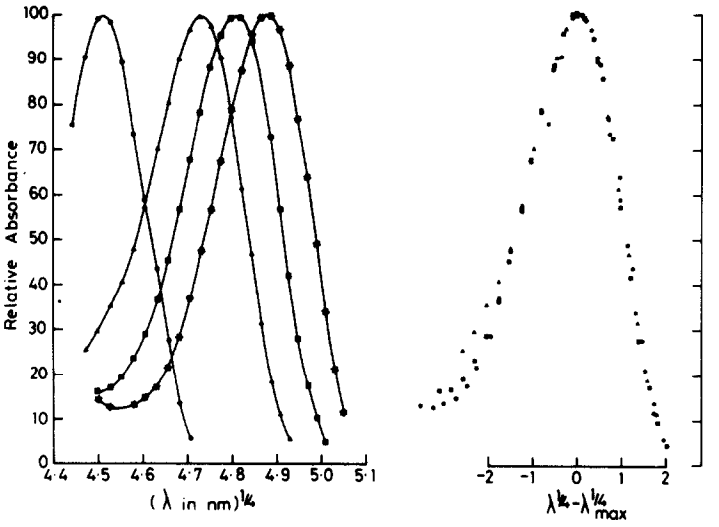

Fig. 6. Relative absorbance of the pigments of the blue cones (●), the rods (▲), the green cones (■) and the red cones (★) of *Macaca fascicularis* [from Bowmaker *et al.* (1980)]. They are plotted on a scale of $(\lambda \text{ in nm})^{1/4}$ at the left, and superimposed on each other on a scale of $\lambda^{1/4}-\lambda^{1/4}_{max}$ at the right. The agreement on the long-wave side is good. On the short-wave side the blue pigment is narrower even near the peak (possibly because of yellow impurities) and the rod pigment is broader lower down the short-wave limb.

tion spectrum is plotted on a scale $f(\lambda)$ instead of λ, it is expanded by a factor $[df(\lambda)/d\lambda]$. One can derive from this the prediction that, if the curves are similar on the $\lambda^{1/4}$ scale, bandwidth (in nm) should vary as the $\frac{3}{4}$ power of wavelength of λ_{max} (in nm). In Fig. 6 data of Ebrey and Honig have been plotted exactly as they did in order to facilitate detailed comparisons, but on double logarithmic scales. For this plot a slope of $-1\frac{1}{4}$ is predicted. There is considerable scatter, as is to be expected with results that were obtained by a variety of methods, but lines of slope $-1\frac{1}{4}$ fit the results for both A_1 and A_2 pigments reasonably well, thus confirming the $\lambda^{1/4}$ rule for both classes of pigment.

TRICHROMACY AND TETRACHROMACY

Figure 5 showed that the blue system would respond to considerably higher comb-frequencies than the green or red and it was the search for a transform of the wavelength scale that would make the frequency responses of the three systems more nearly equal that led to the surprisingly effective $\lambda^{1/4}$ rule for equalizing the shapes of the absorption spectra of pigments. Unfortunately, it is not so effective with the putative sensitivity curves of the human colour mechanisms shown in Fig. 3. When these curves are plotted on a scale of $\lambda^{1/4}$ that for the blue mechanism is narrower than that for the red by some 36% and Fourier transforms correspondingly show that it will pass higher comb-frequencies. From comparison of the spectral sensitivity curves with absorption spectra it is clear that the narrowing of the blue mechanism occurs on the short wavelength side and results from the blue absorption in the pre-receptoral media.

The use of $1/\lambda$ was introduced for pigments, but ironically, although this transform of the wavelength scale actually increases the differences between the pigments, it works rather well in equalizing the Four-

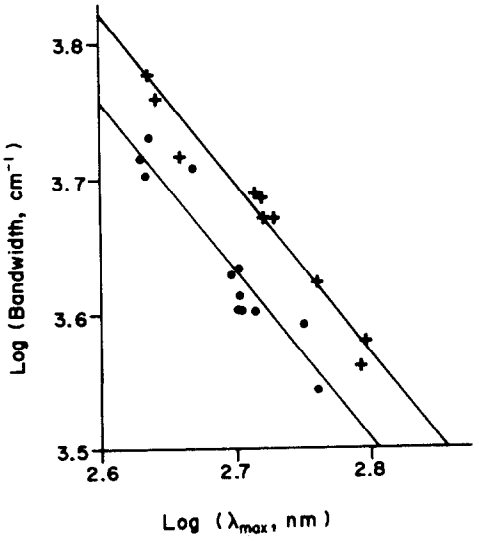

Fig. 7. Ebrey and Honig (1977) assembled the results for many Vit A_1 and Vit A_2 pigments and showed that bandwidth (which they measured in cm^{-1}) increased fairly regularly with the wavelength of the peak absorption (measured in nm). If bandwidth is constant on a $\lambda^{1/4}$ scale, one predicts that bandwidth (in cm^{-1}) will vary as the $-1\frac{1}{4}$ power of λ_{max} (in nm). Their results have therefore been plotted here on double logarithmic co-ordinates and fitted by lines of slope $-5/4$. A_2 pigments (+) are broader than A_1 pigments (●), but as far as one can tell with this amount of scatter the lines are the right slope, thus confirming the $\lambda^{1/4}$ rule.

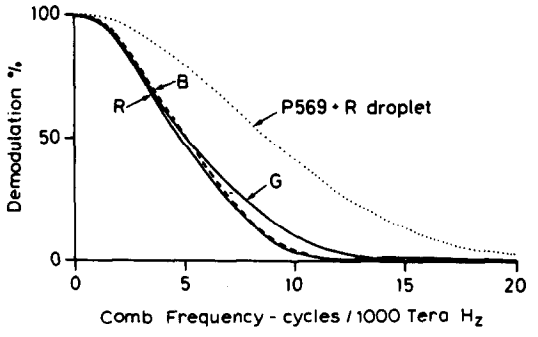

Fig. 8. The three continuous curves are the Fourier transforms (amplitude) of the spectral sensitivities of the three human colour mechanisms [from Smith and Pokorny, (1975)]. These represent the transfer characteristics of the three colour mechanisms for comb-filtered spectra of different comb-frequencies, here expressed in cycles 1000 THz change of spectral frequency. On this scale one cycle of a spectrum of comb-frequency 4.4 cycle/1000 THz would span 435–650 nm. This frequency is demodulated almost to 50% by the human mechanisms, and the comb-frequencies that would become useful if a fourth independent mechanism were present would be more strongly demodulated. The dotted curve is the Fourier transform of the spectral sensitivity of a chicken cone screened by a red oil droplet, as calculated by Bowmaker and Knowles (1977). This responds to much higher comb-frequencies, and tetrachromacy would be perfectly possible without heavy demodulation.

ier transforms of the three human colour mechanisms, as Fig. 8 shows (neglect the dotted curve for the moment). These Fourier transforms represent the responses of the mechanisms for comb-filtered spectra that have uniform wavelength on the $1/\lambda$ scale; the wavelength of such a stimulus lengthens towards the red on a scale of λ, but this slight difference compared with the stimuli of Fig. 3 need not concern us and we can discuss the implications for colour vision of the reduced responses at high comb-frequencies.

The unit for the comb-frequency scale in Fig. 8 is cycles of colour grating per 1000 THz change of light frequency. Taking the limits of wavelength discrimination, as before from Wright and Pitt (1935), at 435 nm (689 THz) and 650 nm (461 THz), one cycle spanning this range (228 THz) has a frequency of 4.4 cycles/1000 THz, and at this frequency the human red, green and blue systems attenuate to 55, 58 and 59%, respectively. Not even a single cycle of a comb-filter can be fitted into the spectrum without serious demodulation, and this result raises an interesting possibility.

Suppose for the sake of argument that instead of attenuating to about 25%, the curve of Fig. 8 had gone right down to zero at 6 cycle/1000 THz; then none of the three colour mechanisms would respond at all to the higher comb-frequencies present in distributions of spectral energy presented to the eye. Now from the sampling theorem (Woodward, 1953; Bracewell, 1965) we know that the number N of independent samples possible in a segment of waveform of extent E and maximum frequency F is $N = 1 + 2EF$. If $E = 167$ THz and F is 6 cycle/1000 THz, this yields $N = 3$ exactly. In other words, it would not be possible to get more than three independent colour variables within this range of the spectrum.* It is usually said that trichromacy results from the fact that there are only three cone pigments, but we now see that quite a different explanation is possible. This might be termed the "low-pass filter theory of trichromacy", and according to it the dimensionality of colour space is determined by the highest comb-frequency passed by the broad spectral sensitivity curves of the photosensitive pigments: even if there were many different types, provided they were all equally broad the representation of a colour would still only require three variables.

In fact, the true figures enable this theory to be rejected. For a cut-off of 8 cycle/1000 THz and a range of 228 THz $N = 4.6$ and of course it becomes even higher if the extended tails of Fig. 8 are included. The highest comb-frequency that three variables suffice to sample adequately in a range of 228 THz is 4.4 cycle/1000 THz; notice, however, that even at this low comb-frequency the breadth of the spectral sensitivity curves causes demodulation almost to 50%, as we started this discussion by pointing out.

The conclusion can be put as follows: another cone pigment could increase the dimensionality of colour space from three to four, and this would in principle allow reconstruction of spectral energy distributions up to comb-frequencies of 6.6 cycle/1000 THz; however, at this comb-frequency the low-pass characteristic of the broad spectral sensitivity curves would attenuate the modulation to 25%. Perhaps it can be concluded that the number and spacing of different photosensitive pigments is adapted to the range of comb-frequencies passed by each one, just as the number and spacing of foveal photoreceptors is adapted in some species to the highest spatial frequencies in the image.

One naturally goes on to ask if there is any fundamental factor that causes spectral sensitivity curves to have the breadth they do. Could visual pigments be evolved with narrower (or broader) absorption spectra? Not nearly enough seems to be known about the theory of absorption spectra to answer this question, but it is suggestive to look now at the dotted curve of Fig. 8. This is the Fourier transform of the spectral sensitivity of cones containing the 569-nm λ_{max} pigment of chickens together with the red oil droplet, calculated by Bowmaker and Knowles (1977). This extends to comb-frequencies 70% higher than the three human colour mechanisms. One immediately obvious consequence of this is that tetrachromacy

*It is, strictly speaking, impossible for a finite segment of waveform to have a finite frequency limit and vice versa. The value of N can be taken to suggest a practical limit to the number of independent samples; signal–noise considerations would rapidly decrease the information available from additional samples.

would be possible for chickens with greatly reduced losses from demodulation of high comb-frequencies. The selective advantage of oil droplets is not altogether clear on other grounds, and this narrowing of the spectral sensitivity curves, making higher dimensionality of the colour space useful, may be the answer.

PHASE DISCRIMINATION

The original discussion with John Mollon which prompted this article arose from his doubting if broad spectral sensitivity curves were as bad for colour discrimination as I had suggested. I think it is clear from the narrower Fourier transforms of the broad spectral sensitivity curves that the latter are bound to be worse at discriminating the shift of a spectral line. For a single pigment, the shift of a spectral line is exactly equivalent to a change of phase of all the comb-frequencies of which the line is composed and the phase change (in radians) is directly proportional to comb-frequency. Hence attenuating the higher comb-frequencies removes the very components which change most. There are, however, mitigating factors that make broad bands a better choice. First, they will collect more light, other things being equal, and this will improve their sensitivity. Second, the lights that have to be discriminated under normal conditions are not like monochromatic bands, for they result from relatively broad-band natural pigments. Since they will not contain high comb-frequencies, the low-pass characteristic of the colour mechanisms will have little effect on their discriminability.

POSITIONS OF THE PEAKS OF RED, GREEN AND BLUE SYSTEMS

So far it has been suggested that the useful number of different colour mechanisms is limited to three (in mammals) because the broad spectral sensitivity curves act as low-pass filters demodulating the high comb-frequencies that would make it advantageous to have more than three independent dimensions. Nothing has been said about the positioning of the peaks of the individual mechanisms, but one might expect them to be equally spaced, one at each extreme and one in the middle, in order to obtain maximum independence of each other. The actual arrangement is very different, however. Taking 435 and 650 nm as the limits, it is a surprise to find the peaks at 440, 545 and 570 nm; over 100 nm separate blue from green and only 25 nm the green from the red. Expressing these as spectral frequencies, the limits are 689 and 461 THz, and the peaks are at 681 THz for the blue, and 550 and 526 THz for the other two. Hence the unequal spacing is even more pronounced.

The result of this arrangement is to make the activation of red and green systems strongly correlated, as is indeed obvious from the overlap of their spectral sensitivity curves. Though they have a degree of independence which enables them to signal two dimensions in the colour space, they will cover this like measurements in a two-dimensional plane along two axes that are at a narrow angle to each other. The only possible merit of this arrangement that I can suggest is that, because the two curves are so close together, they can be treated alike for problems of spatial resolution. I think it is usually supposed that red and green systems each have a capacity for spatial resolution adequate for the quality of the optical image on the retina. According to the current suggestion there is only one mechanism of high spatial resoltuion and it is fed by both red and green cones without distinction; differences of excitation caused by colour are ignored and could only be detrimental to spatial resolution. I have not found any data that definitely rule out this possibility, though one might predict small differences in acuity dependent on wavelength that have not been reported, and dichromats and anomalous trichomats might be expected to have improved acuity under some conditions compared with normals.

A very large number of people must have discussed a problem with William Rushton, and very few can have failed to gain new insight as a result. Why was this so? As well as being inquisitive and highly rational he was also a very colourful personality, and I think it was this aspect—his gift for presenting a problem dramatically and from an unfamiliar angle—that enabled him to add new dimensions to one's understanding. I cannot match his drama, but this paper was written in the spirit or a discussion with William, and I hope it may have given some readers an idea how his unfamiliar viewpoints could enlarge one's understanding.

Acknowledgements—I am much indebted to J. Bowmaker, H. Dartnall and J. Mollon for giving me some of their latest results from microspectrophotometry of photoreceptors, and to the latter for provoking the arguments presented here. Some earlier work on similar lines was done while working on NIHS grant EY3412 to Dr Sakitt.

REFERENCES

Barlow H. B. (1965) Visual resolution and the diffraction limit. *Science* **149**, 553–555.

Barlow H. B. and Mollon J. D. (Eds) (1982) *The Senses*. Cambridge University Press, Cambridge.

Bowmaker J. K., Dartnall H. J. A. and Mollon J. D. (1980) Microspectrophotometric demonstration of four classes of photoreceptor in an Old World primate *Macaca fascicularis*. *J. Physiol.* **298**, 131–143.

Bowmaker J. K. and Knowles A. (1977) The visual pigments and oil droplets of the chicken retina. *Vision Res.* **17**, 755–764.

Boynton R. M. (1979) *Human Color Vision*. Holt, Rinehart & Winston, New York.

Bracewell R. (1965) *The Fourier Transform and its Applications*. McGraw-Hill, New York.

Brindley G. S. (1955) The colour of light of very long wavelength. *J. Physiol.* **130**, 35–44.

Campbell F. W. and Robson J. G. (1968) Applications of Fourier analysis to the visibility of gratings. *J. Physiol.* **197**, 551–566.

Dartnall H. J. A. (1953) The interpretation of spectral sensitivity curves. *Br. med. Bull.* **9**, 24–30.

Ebrey T. G. and Honig B. (1977) New wavelength-dependent visual pigment monograms. *Vision Res.* **17**, 147–151.

Enroth-Cugell C. and Robson J. G. (1966) The contrast sensitivity of retinal ganglion cells of the cat. *J. Physiol.* **187**, 517–552.

de Lange H. (1952) Experiments on flicker and some calculations on an electrical analogue of the foveal systems. *Physica* **18**, 935–950.

Lewis P. R. (1955) A theoretical interpretation of spectral sensitivity curves at long wavelengths. *J. Physiol.* **130**, 45–52.

Schade O. H. (1956) Optical and photoelectric analog of the eye. *J. opt. Soc. Am.* **46**, 721–739.

Smith V. C. and Pokorny J. (1975) Spectral sensitivity of the foveal cone photopigments between 400 and 500 nm. *Vision Res.* **15**, 161–171.

Stiles W. S. (1948) The physical interpretation of the spectral sensitivity curve of the eye. *Worshipful Company of Spectacle Makers Convention*, Nov. 1948.

Woodward P. M. (1953) *Probability and Information Theory with Applications to Radar.* Pergamon Press, Oxford.

Wright W. D. and Pitt F. H. G. (1935) The colour vision characteristics of two trichromats. *Proc. physiol. Soc., Lond.* **47**, 207–208.