FUNDAMENTALS OF COLOUR TELEVISION

by F. W. de VRIJER.

621.397.2

In recent years the colour printing of illustrations in magazines and books, and the colour film in the cinema and in amateur photography, have attained a high degree of perfection. They have, as it were, opened the general public's eyes to the world of colour. Colour television has also made considerable strides, as evidenced, for example, by the regular colour television transmissions in the United States. In view of the high costs at present involved, colour transmissions are being made on a relatively small scale; it is not to be expected, therefore, that colour television will, for the time being, be introduced in Europe other than experimentally.

The article below is an introduction to a series of articles to be printed in this Review describing some of the work being done by Philips in the field of colour television.

One of the fundamentals of colour television is, of course, the science of colour. Various colorimetric concepts, such as the chromaticity diagram (colour triangle), have already been discussed in this Review [1]). We shall begin by briefly recapitulating these concepts.

Colorimetry

The spectral energy distribution of a light source determines the colour of the light. On the other hand, the colour of the light does not uniquely establish the spectral distribution, for different spectral distributions can produce the same colour sensation. It was known to Newton that almost all colours can be imitated by additively mixing light of only three primary colours, i.e. by so combining them that the various types of light impinge simultaneously on the same part of the observer's retina. Additive mixing occurs, for example, when two or more projectors, each with a light beam of a given colour, illuminate the same part of a projection screen. If red, green and blue are taken as the primary colours, almost all colours can be reproduced by additive mixing (the minor reservation contained in the word "almost" will be explained later). By adding green to red, for instance, we get orange; a little more green produces yellow and still more green results in yellowish green. Mixed in the right proportions, the three primary colours can produce white (*fig. 1*).

The converse to additive mixing is subtractive mixing. This arises when specific parts of the spectrum of a light source are attenuated more than others, for example by passing the light through a filter or by making it reflect selectively from a surface. The latter is involved in the mixing of pigments [2]) and is, for example, the basis of modern forms of colour photography. The laws of subtractive mixing are more complicated than those of additive mixing (we shall return to this subject later). *Fig. 2* gives an example of the difference in result as against fig. 1, and shows the effects of mixing two and three paints whose colours are complementary to those of the light on which fig. 1 is based.

The primary colours frequently used in colorimetry are the spectral colours given below and defined by their associated wavelengths λ (in vacuo):

$$\left.\begin{array}{ll}\text{Red,} & \lambda = 700.0\ \mathrm{m\mu} \\ \text{Green,} & \lambda = 546.1\ \mathrm{m\mu} \\ \text{Blue,} & \lambda = 435.8\ \mathrm{m\mu}\end{array}\right\} \quad . \; . \; . \quad (1)$$

If the eye is unable to distinguish a given type of light from the additive mixture of B_1 lumens at $\lambda = 700.0\ \mathrm{m\mu}$, B_2 lumens at $\lambda = 546.1\ \mathrm{m\mu}$ and B_3 lumens at $\lambda = 435.8\ \mathrm{m\mu}$, we then say that the light concerned has the colour coordinates B_1, B_2 and B_3 in the system based on the above-mentioned primary colours. *Fig. 3* gives the result of measurements — averaged over a large number of observers — of the primary-colour coordinates of 1 lumen of any spectral colour. It can be seen that negative values also occur. For example, if B_1 is negative while B_2 and B_3 are positive, it means that the additive mixture of the original light and $(-B_1)$ lumens of $\lambda = 700.0\ \mathrm{m\mu}$, appears the same to the eye as the additive mixture of B_2 lumens of $\lambda = 546.1\ \mathrm{m\mu}$ and B_3 lumens of $\lambda = 435.8\ \mathrm{m\mu}$. Such types of light, with one or two negative coordinates, cannot therefore be imitated by additively mixing the three selected primary colours.

[1]) See W. de Groot and A. A. Kruithof, Philips tech. Rev. **12**, 137-144, 1950/51. For a fuller treatment see P. J. Bouma, Physical aspects of colour (Philips Technical Library 1947).

[2]) See J. L. H. Jonker and S. Gradstein, Fluorescent pigments as an artistic medium, Philips tech. Rev. **11**, 16-22, 1949/50.

If, instead of one spectral colour, we have light with a given spectral energy distribution $E(\lambda)$, we may write for the coordinates:

$$\left.\begin{aligned} B_1 &= \int \overline{B}_1(\lambda)E(\lambda)\,\mathrm{d}\lambda, \\ B_2 &= \int \overline{B}_2(\lambda)E(\lambda)\,\mathrm{d}\lambda, \\ B_3 &= \int \overline{B}_3(\lambda)E(\lambda)\,\mathrm{d}\lambda. \end{aligned}\right\} \quad . \; . \; . \; . \quad (2)$$

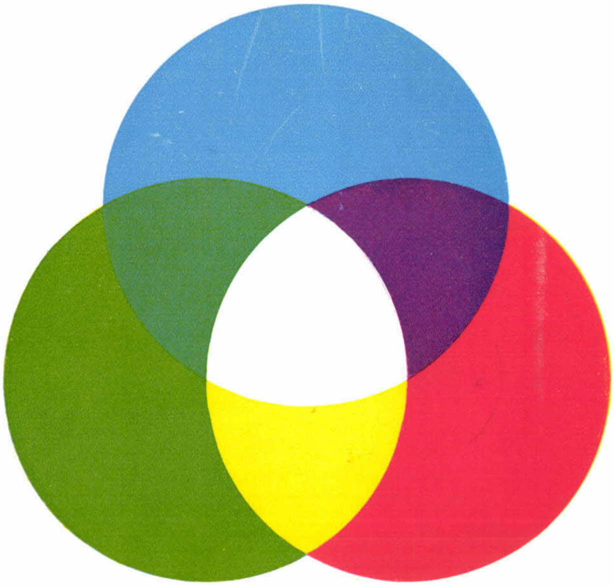

Fig. 1.

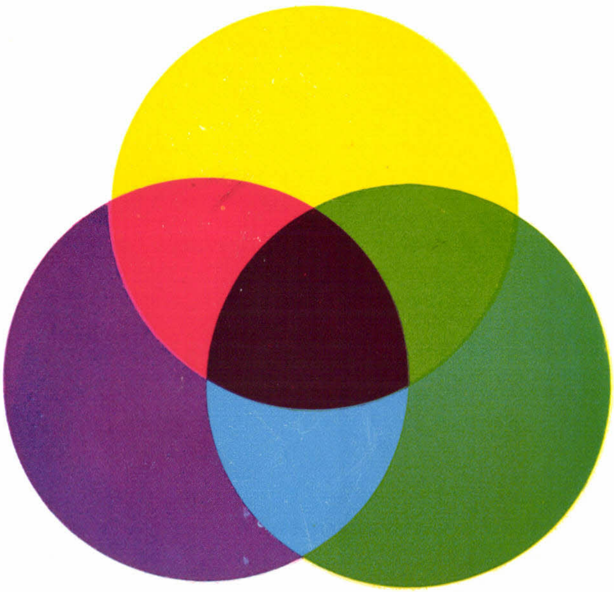

Fig. 2.

Fig. 1. Additive mixing of red, green and blue light. Mixed by twos, these types of light produce yellow, blue-green or purple (magenta). The three types of light mixed in the right proportions produce white.
Fig. 2. Subtractive mixing. A yellow dye mixed with a blue-green dye produces a green mixture, a blue-green and a purple dye produce blue, and a purple and a yellow dye produce red. The three dyes together can produce black.

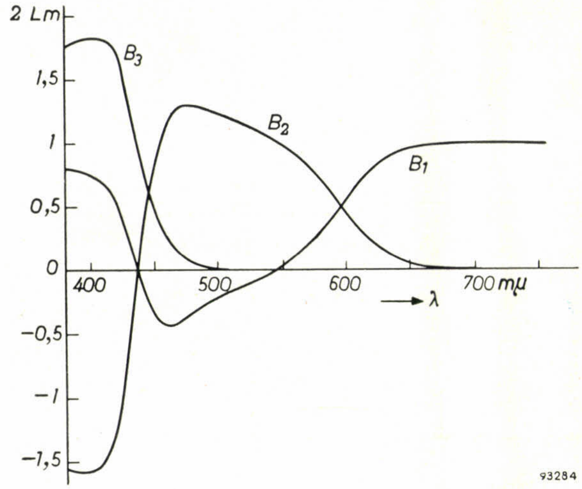

Fig. 3. The colour coordinates B_1, B_2 and B_3 of 1 lumen of the spectral colours with wavelength λ, in the system of primary colours given by (1). (1 m$\mu = 10^{-9}$ m.)

The quantities $\overline{B}_1(\lambda)$, $\overline{B}_2(\lambda)$ and $\overline{B}_3(\lambda)$ are represented graphically in *fig. 4*.

It appears, then, that the colour coordinates of a mixture of spectral colours are additive. This additivity applies not only to spectral colours but, as will be clear, it also applies generally to arbitrary spectral distributions.

Two types of light with the same ratio $B_1 : B_2 : B_3$ have the same "type" of colour (or *chromaticity*), but in general they differ in luminance. Luminance has been defined by the Commission Internationale de l'Eclairage (C.I.E.) as the luminous flux

$$K \int V(\lambda)\, E(\lambda)\, \mathrm{d}\lambda \quad . \; . \; . \; . \; . \; . \quad (3)$$

per unit area and per unit solid angle. In the expression (3) the constant K is approximately 683 lumens/watt, and $V(\lambda)$ is the internationally accepted relative spectral sensitivity of the eye, determined from experiments using large numbers of observers. "Relative" implies here that the maximum of $V(\lambda)$ is made equal to 1; for wavelengths

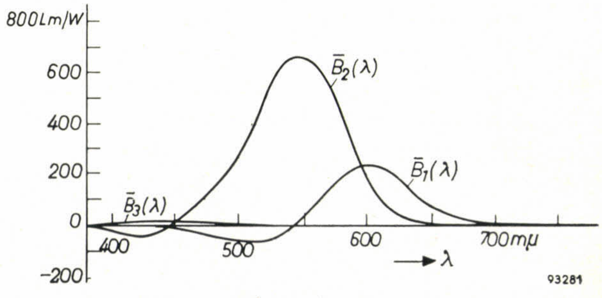

Fig. 4. The quantities $\overline{B}_1(\lambda)$, $\overline{B}_2(\lambda)$ and $\overline{B}_3(\lambda)$ from (2) as a function of λ. This figure differs from fig. 3 in that it is based not on the luminous flux but on the energy flux E. Fig. 4 can be derived from fig. 3 by multiplication by the visual-sensitivity curve.

smaller than 400 mμ and larger than 780 mμ [3]), $V(\lambda)$ is practically zero. The unit of luminance is the candela per m^2.

From the additivity rule it follows directly that the luminance L of any type of light is

$$L = B_1 + B_2 + B_3 = \int \{\overline{B}_1(\lambda) + \overline{B}_2(\lambda) + \overline{B}_3(\lambda)\} E(\lambda) \mathrm{d}\lambda.$$

Thus, according to (3),

$$KV(\lambda) = \overline{B}_1(\lambda) + \overline{B}_2(\lambda) + \overline{B}_3(\lambda). \quad . \quad . \quad (4)$$

Instead of taking the coordinates B_1, B_2 and B_3 to characterize a type of light, we may also take three mutually independent linear combinations of B_1, B_2 and B_3. The C.I.E. has laid down as the international system of colour coordinates:

$$\left.\begin{aligned} X &= 2.7689\, B_1 + 0.38159\, B_2 + 18.801\, B_3, \\ Y &= B_1 + B_2 + B_3, \\ Z &= 0.01237\, B_2 + 93.060\, B_3. \end{aligned}\right\} \quad . \quad (5)$$

From this it follows that:

$$\left.\begin{aligned} X &= K \int \overline{X}(\lambda)\, E(\lambda)\, \mathrm{d}\lambda, \\ Y &= K \int \overline{Y}(\lambda)\, E(\lambda)\, \mathrm{d}\lambda, \\ Z &= K \int \overline{Z}(\lambda)\, E(\lambda)\, \mathrm{d}\lambda, \end{aligned}\right\} \quad . \quad . \quad . \quad (6)$$

where

$$\begin{aligned} K\overline{X}(\lambda) &= 2.7689\overline{B}_1(\lambda) + 0.38159\overline{B}_2(\lambda) + 18.801\overline{B}_3(\lambda), \\ K\overline{Y}(\lambda) &= \overline{B}_1(\lambda) + \overline{B}_2(\lambda) + \overline{B}_3(\lambda), \\ K\overline{Z}(\lambda) &= 0.01237\overline{B}_2(\lambda) + 93.060\overline{B}_3(\lambda). \end{aligned}$$

(See in this connection page 137 of the article quoted in footnote [1].) A graphic representation of $\overline{X}(\lambda)$, $\overline{Y}(\lambda)$ and $\overline{Z}(\lambda)$ is given in *fig. 5*. Two advantages of this X, Y, Z system [4]) are:

1) one of the coordinates, viz. Y, is identical with the luminance;
2) negative values of the coordinates do not occur.

With reference to the latter advantage, it may be remarked that not every set of X, Y, Z values, although none are negative, corresponds to an existing type of light. For instance, colours corresponding to points on the coordinate axes (two of the three coordinates being zero) do not exist. This is sometimes expressed by saying that "the primary colours of the X, Y, Z system are not real".

It has already been pointed out that types of light with the same ratio of coordinates have the same chromaticity (type of colour) and merely differ in luminance. Clearly this also applies to the X, Y, Z system. To designate the colour, therefore, two quantities are sufficient. Those frequently used are:

$$x = \frac{X}{X+Y+Z} \text{ and } y = \frac{Y}{X+Y+Z}. \quad . \quad . \quad (7)$$

Every existing colour may now be represented by a point in a flat plane, with the coordinates x and y. All these points together fill what is known as the chromaticity diagram or colour triangle (*fig. 6*) [5]).

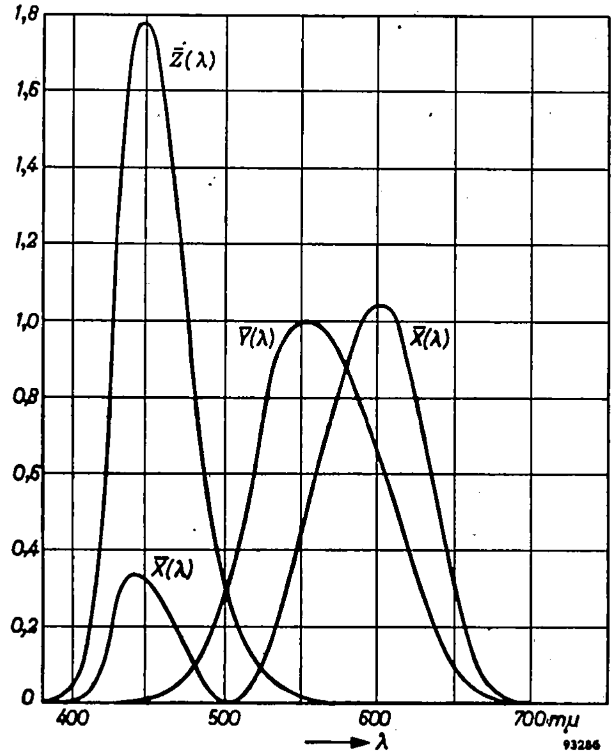

Fig. 5. The quantities $\overline{X}(\lambda)$, $\overline{Y}(\lambda)$ and $\overline{Z}(\lambda)$ from (6) as a function of λ.

The periphery is occupied by the spectral colours and the so-called purple or magenta colours, the middle by white. From the periphery towards white the "saturation" decreases from maximum to zero. The caption of fig. 6 explains the rule for additive mixing (centre-of-gravity rule).

Principles of colour television

Starting from the colorimetric principles described above, we can imagine a colour television system as being realized in the manner shown schematically in *fig. 7*. The light L coming from the scene is split into three components, red, green and blue. These three components each act upon a camera tube (E_r, E_g, E_b), the video signals of which are transmitted via separate cables or radio channels to the associated picture tubes (F_r, F_g, F_b) at the reception

[3]) In addition to the objective concept "luminance" there exists the subjective concept of "brightness". Although brightness is almost synonymous with luminance, it is better not to use the word brightness in a quantitative sense.

[4]) For other advantages (not of importance in colour television) see Chapters 5 and 6 of the book by Bouma mentioned in [1]).

[5]) A reproduction in colour is given in Philips tech. Rev. 12, 141, 1950/51.

end. The latter tubes are provided respectively with a red, a green and a blue luminescent phosphor. Combined in the right way, the three monochromatic pictures produce a single multicoloured picture. (Instead of the projection system as shown here at the receiving end, which is more suitable for the production of large pictures, use is preferably made for home-reception of a special direct-view tube; we shall not discuss this system here.) For the sake of clarity, the amplifiers required for boosting the signal to the picture tubes, and the deflection and synchronization devices, are not shown in fig. 7.

The separation of the incoming light into three components at the transmitting end, as well as the combination of the three pictures at the receiving end, can be effected by means of dichroic mirrors, i.e. mirrors which efficiently reflect a part of the spectrum and transmit the remaining part. An article has recently appeared in this Review, describing the action of these mirrors, which is based on interference, and discussing their manufacture by the deposition of thin layers [6]). This article also describes the principle of combining the three pictures: the red reflecting mirror M_{r2} (fig. 7) transmits green and blue, the blue reflecting mirror M_{b2} transmits green and red, and with their aid the three differently coloured pictures on the receiving tubes (in this case projection tubes) can be projected on to a screen S in accurate register and with little loss of light. This projection system will be dealt with in more detail in a subsequent article.

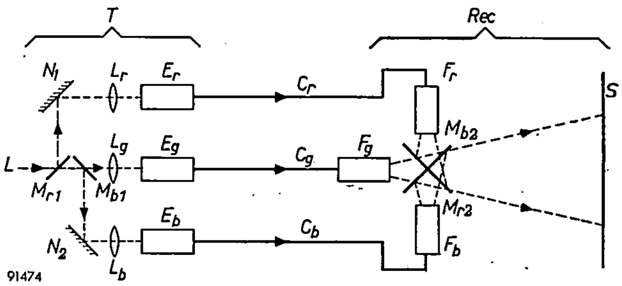

Fig. 7. Schematic representation of a system for colour television.

At the transmission end T, of the light L coming from the scene, the red reflecting mirror M_{r1} separates the red component, which is projected on to the "red" camera tube E_r via the ordinary mirror N_1 and the lens L_r. The blue component reaches the "blue" camera tube E_b via the blue reflecting mirror M_{b1}, the ordinary mirror N_2 and the lens L_b. The green component passes through M_{r1} and M_{b1} and arrives at the "green" camera tube E_g via the lens L_g. Separate cables or radio links C_r, C_g and C_b convey the "red", "green" and "blue" signals to the receiving end *Rec*, where they produce a picture of corresponding colour on picture tubes F_r, F_g and F_b with red, green and blue luminescent phosphors respectively. These picture tubes are each associated with a projector. By means of the dichroic mirrors M_{r2} and M_{b2}, the three pictures are projected in exact register on to the screen S. (Since the beams have a large cross-section in this case, intersecting dichroic mirrors are often used; see [6]).)

Fig. 7 also shows the way in which the incoming light is split into three components, by means of the dichroic mirrors M_{r1} and M_{b1} and the ordinary mirrors N_1 and N_2. This needs no further explanation. It may be added that colour filters can be placed in the three light paths in order to produce a closer approximation to the required spectral sensitivities; we shall return to this later.

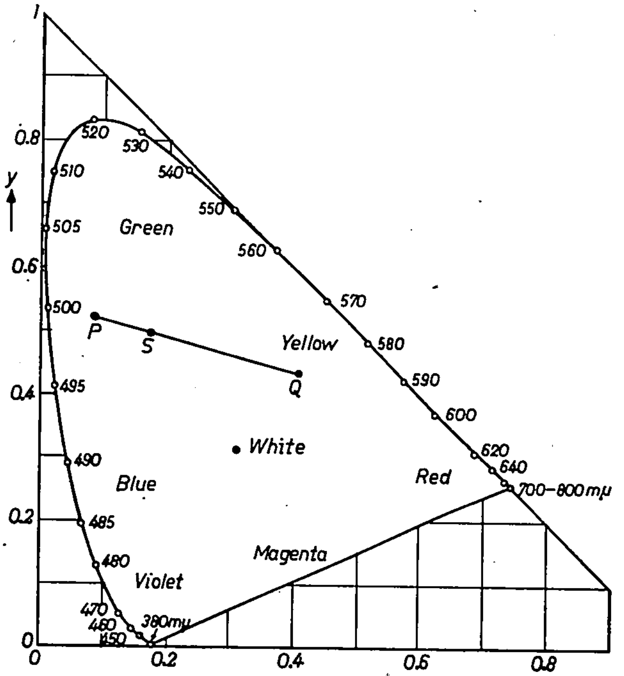

Fig. 6. The colour triangle [5]). The wavelengths of the spectral colours are given around the circumference.

Rule for additive mixing. Assume that Y_P lumens of colour P (with coordinates x_P, y_P in the colour triangle) are mixed with Y_Q lumens of the colour Q (x_Q, y_Q). The international colour coordinates X, Y and Z are then:

	X	Y	Z
Colour P	$X_P = x_P Y_P/y_P$	Y_P	$Z_P = (1-x_P-y_P)\,Y_P/y_P$
Colour Q	$X_Q = x_Q Y_Q/y_Q$	Y_Q	$Z_Q = (1-x_Q-y_Q)\,Y_Q/y_Q$

The coordinates of the colour mixture are found by addition. The result of the mixture has a luminance $Y_P + Y_Q$ and its colour point (S) follows from the centre-of-gravity rule. Imagine that the mass $X_P + Y_P + Z_P$ is fixed at P and the mass $X_Q + Y_Q + Z_Q$ at Q. The centre of gravity S of these masses is then the required colour point, which always lies on the connecting line PQ. Since all types of light consist of a discrete line spectrum or a continuous spectrum, the colour points of all existing types of light lie in the area bounded by the line of the spectral colours and the so-called purple line.

On the basis of this extremely simplified colour television system we shall now consider various questions more in detail.

Choice of the primary colours

A colour produced by the additive mixing of two other colours always lies in the colour triangle upon the line connecting these two colours, at a position

[6]) P. M. van Alphen, Philips tech. Rev. 19, 55-63, 1957/58 (No. 2).

which follows from the centre-of-gravity rule (fig. 6). Starting from three primary colours we can therefore reproduce only the colours that lie within (or on the circumference of) the triangle formed by the colour points of the three selected primary colours. This triangle must accordingly be large enough to contain that part of the colour triangle which is of practical importance. It follows directly from the position of the colours in the colour triangle that, broadly speaking, red, green and blue are the appropriate primary colours.

In order to determine more exactly the primary colours required for a television system it is necessary to take account of some important practical considerations. At the receiving end it is desirable to be able to produce the three primary colours by means of known phosphors, and to do this with an efficiency capable of giving the required luminance without having to apply too great a power to the picture tubes. In the United States these considerations have led to the following choice of primary colours:

$$\left.\begin{array}{lcc} & x & y \\ \text{Red} \ \ldots & 0.67 & 0.33 \\ \text{Green} \ \ldots & 0.21 & 0.71 \\ \text{Blue} \ \ldots & 0.14 & 0.08 \end{array}\right\} \quad \ldots \quad (8)$$

In *fig. 8* these colour points (which differ from the spectral primary colours given by (1)) are designated by R, G and B. With these primary colours, then, it is possible to produce all colours whose colour point lies within (or on) the triangle RGB. It has been found that almost all the colours occurring in nature, and also those of common dyes (bounded by contour *1* of fig. 8) fall within this triangle. Compared with other well-known colour reproduction processes (colour photography, for example) the part of the colour triangle covered in this case may be regarded as favourable (compare contour *2* in fig. 8 with the triangle RGB). In the following pages we shall base our discussion on the primary colours given by (8).

Reproduction of white

The reproduction of white calls for special attention. The colour point of white, as radiated by a picture tube for monochromatic television, generally lies in the region of so-called standard white C (point C in fig. 8), with the coordinates $x = 0.310$, $y = 0.316$. This colour is generally accepted as a good choice, and it is reasonable, therefore, to take it as the white for colour television too, i.e. a colourless object in the scene should be reproduced by standard white C. A simple calculation shows that in order to produce 1 candela/m² of this white light from the three primary colours (8), we need 0.30 cd/m² red, 0.59 cd/m² green and 0.11 cd/m² blue.

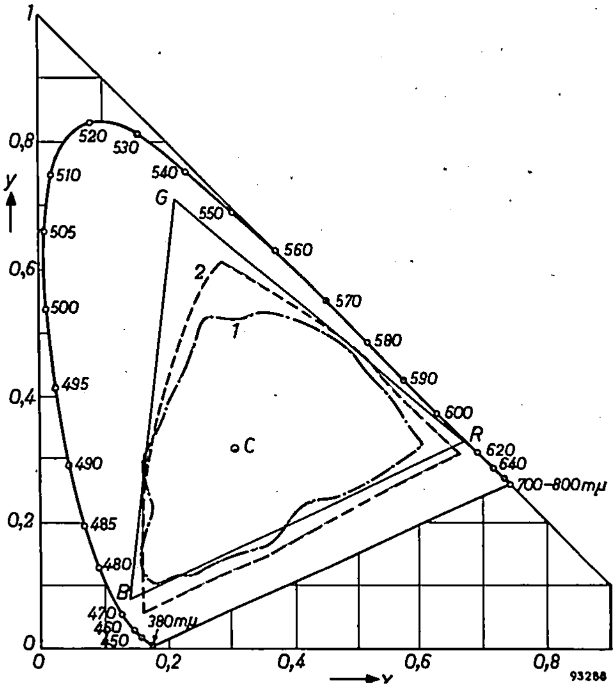

Fig. 8. Colour triangle, with the colour point C of standard white C ($x = 0.310$, $y = 0.316$). R, G and B are the primary colours selected for colour television, with coordinates according to (8). By additive mixing, all colours can be produced whose colour point lies within or on the triangle RGB. This is more favourable than the possibilities in subtractive colour-film processes, in which the contour *2* limits the realizable colours. Within contour *1* lie all the reflection colours occurring in nature and the reflection colours of all known dyes and printing inks.

This calculation is made as follows. For 1 cd/m² white light with $x = 0.310$, $y = 0.316$, $z = 1-(x+y) = 0.374$, the associated values of X, Y and Z are:

$$X_w = \frac{x}{y} = \frac{0.310}{0.316} = 0.980,$$
$$Y_w = 1.000,$$
$$Z_w = \frac{z}{x} X_w = \frac{0.374}{0.310} \times 0.980 = 1.184.$$

l_r cd/m² red with $x = 0.67$, $y = 0.33$, $z = 0$ gives:

$$\left.\begin{aligned} X_r &= \frac{0.67}{0.33} l_r = 2.03\, l_r, \\ Y_r &= 0.33\, l_r, \\ Z_r &= 0. \end{aligned}\right\} \quad \ldots\ldots\ldots \quad (9a)$$

l_g cd/m² green with $x = 0.21$, $y = 0.71$, $z = 0.08$ gives

$$\left.\begin{aligned} X_g &= \frac{0.21}{0.71} l_g = 0.296\, l_g, \\ Y_g &= l_g, \\ Z_g &= \frac{0.08}{0.71} l_g = 0.113\, l_g. \end{aligned}\right\} \quad \ldots\ldots\ldots \quad (9b)$$

l_b cd/m² blue with $x = 0.14$, $y = 0.08$, $z = 0.78$ gives

$$\left.\begin{aligned} X_b &= \frac{0.14}{0.08} l_b = 1.75\, l_b, \\ Y_b &= l_b, \\ Z_b &= \frac{0.78}{0.08} l_b = 9.75\, l_b. \end{aligned}\right\} \quad \ldots\ldots\ldots \quad (9c)$$

In order to obtain exactly 1 cd/m² standard white C by additively mixing l_r cd/m² of this red, l_g cd/m² of this green and l_b cd/m² of this blue, we must ensure that the following relations exist:

$$X_w = 2.03\ l_r + 0.296\ l_g + 1.75\ l_b = 0.980\,,$$
$$Y_w = l_r + l_g + l_b = 1.000\,,$$
$$Z_w = 0.113\ l_g + 9.75\ l_b = 1.184\,.$$

From these three equations we can solve l_r, l_g and l_b. We find: $l_r = 0.298$ cd/m², $l_g = 0.588$ cd/m², $l_b = 0.114$ cd/m² (rounded off: 0.30, 0.59 and 0.11 cd/m²).

Luminance signal

In colour television the signal amplitude is commonly standardized so as to make the signals R (red), G (green) and B (blue) of equal strength for white, the relation for white thus being:

$$R = G = B\,.$$

Moreover, it is usual to give the value of 1 to the signal values for the white with the greatest luminance occurring in the scene. Thus, for this white of maximum luminance

$$R = G = B = 1\,.$$

If this white is reproduced with a luminance L, the red primary colour contributes 0.30 L, the green 0.59 L, and the blue 0.11 L; the sum is, of course, exactly L. These contributions of luminance are therefore in this case the luminances of the primary colours corresponding to the signal values 1.

For the sake of convenience we shall now assume that the display device functions linearly, i.e. that for each primary colour the reproduced luminous flux is proportional to the signal strength R, G or B. For each value of R, G and B the luminance contributions will then be respectively $0.30\,RL$ (red), $0.59\,GL$ (green) and $0.11\,BL$ (blue). We may then write for each colour that the luminance is equal to:

$$(0.30\,R + 0.59\,G + 0.11\,B)L\,.$$

The linear combination

$$H = 0.30\,R + 0.59\,G + 0.11\,B \quad . \ . \ (10)$$

is called the luminance signal. For white with the maximum luminance ($R = G = B = 1$), the value $H = 1$. The luminance signal corresponds to the normal video signal of monochromatic television.

Still assuming the display device to function linearly, we can calculate as follows the signals R, G and B needed for producing a colour with coordinates X, Y, Z.

According to the normalizing of R, G and B mentioned above,

$R = 1$ corresponds to 0.298 L cd/m² of the primary red,
$G = 1$ to 0.588 L cd/m² of the primary green, and
$B = 1$ to 0.114 L cd/m² of the primary blue.

It now follows directly from (9a), (9b) and (9c) that for given R, G and B the coordinates X, Y and Z of the reproduced colour will satisfy:

$$X = 2.03\times0.298\,RL + 0.296\times0.588\,GL + 1.75\times0.114\,BL,$$
$$Y = 0.298\,RL + 0.588\,GL + 0.114\,BL,$$
$$Z = 0.113\times0.588\,GL + 9.75\times0.114\,BL.$$

For this we may write:

$$\frac{X}{L} = 0.604\,R + 0.174\ G + 0.200\,B,$$
$$\frac{Y}{L} = 0.298\,R + 0.588\ G + 0.114\,B,$$
$$\frac{Z}{L} = 0.000\,R + 0.0664\,G + 1.112\,B.$$

For R, G and B as functions of X, Y and Z we find from this:

$$\left.\begin{aligned} RL &= 1.92\ \ X - 0.535\ Y - 0.290\ \ Z,\\ GL &= -0.984\ \ X + 2.00\ \ Y - 0.0274\ Z,\\ BL &= 0.0588\ X - 0.119\ Y + 0.901\ \ Z.\end{aligned}\right\} \quad . \ . \ (11)$$

In fact, normal display devices do *not* function linearly. This necessitates certain corrections to the above considerations (gamma correction), which we shall discuss at the end of this article.

Signals R, G and B at the transmitting end

We shall now say a few words about the generation of signals R, G and B at the camera end.

If $E(\lambda)$ is the spectral energy distribution of a given colour in the scene, we may write for the signals to be generated:

$$\left.\begin{aligned} R &= \int\overline{R}(\lambda)\ E(\lambda)\ \mathrm{d}\lambda,\\ G &= \int\overline{G}(\lambda)\ E(\lambda)\ \mathrm{d}\lambda,\\ B &= \int\overline{B}(\lambda)\ E(\lambda)\ \mathrm{d}\lambda.\end{aligned}\right\} \quad . \ . \ . \ . \ (12)$$

$\overline{R}(\lambda)$, $\overline{G}(\lambda)$ and $\overline{B}(\lambda)$ are the total spectral sensitivities of the three channels. How must these be chosen in order to obtain exact colour reproduction? The necessary variation of $\overline{R}(\lambda)$, $\overline{G}(\lambda)$ and $\overline{B}(\lambda)$ as a function of λ can be calculated; the result (apart from the constant L/K) is shown in *fig. 9*. Exact colour reproduction is only possible for the colours inside the triangle RGB; only for these colours are none of the signals R, G and B negative.

From (11) and (6) it follows that for exact colour reproduction the signals R, G and B must satisfy:

$$\frac{L}{K}R = 1.92\ \int\overline{X}(\lambda)\,E(\lambda)\,\mathrm{d}\lambda - 0.535\int\overline{Y}(\lambda)\,E(\lambda)\,\mathrm{d}\lambda - 0.290\int\overline{Z}(\lambda)\,E(\lambda)\,\mathrm{d}\lambda,$$
$$\frac{L}{K}G = -0.984\ \int\overline{X}(\lambda)\,E(\lambda)\,\mathrm{d}\lambda + 2.00\int\overline{Y}(\lambda)\,E(\lambda)\,\mathrm{d}\lambda - 0.0274\int\overline{Z}(\lambda)\,E(\lambda)\,\mathrm{d}\lambda,$$
$$\frac{L}{K}B = 0.0588\int\overline{X}(\lambda)\,E(\lambda)\,\mathrm{d}\lambda - 0.119\int\overline{Y}(\lambda)\,E(\lambda)\,\mathrm{d}\lambda + 0.901\int\overline{Z}(\lambda)\,E(\lambda)\,\mathrm{d}\lambda.$$

By comparing these three expressions with (12), it follows directly that:

$$\frac{L}{K}\overline{R}(\lambda) = 1.92\,\overline{X}(\lambda) - 0.535\,\overline{Y}(\lambda) - 0.290\,\overline{Z}(\lambda),$$
$$\frac{L}{K}\overline{G}(\lambda) = -0.984\,\overline{X}(\lambda) + 2.00\,\overline{Y}(\lambda) - 0.0274\,\overline{Z}(\lambda),$$
$$\frac{L}{K}\overline{B}(\lambda) = 0.0588\,\overline{X}(\lambda) - 0.119\,\overline{Y}(\lambda) + 0.901\,\overline{Z}(\lambda).$$

$\overline{X}(\lambda)$, $\overline{Y}(\lambda)$ and $\overline{Z}(\lambda)$ are here known functions of λ, viz. the spectral distribution curves for the "spectrum of constant energy flux", which also occur in (6). See fig. 5.

It can be seen from fig. 9 that a negative sensitivity is required in certain wavebands. Since this requirement cannot easily be put into effect, and since its neglect results in only minor errors in the colour reproduction, the negative portions of the curves are normally left out of consideration.

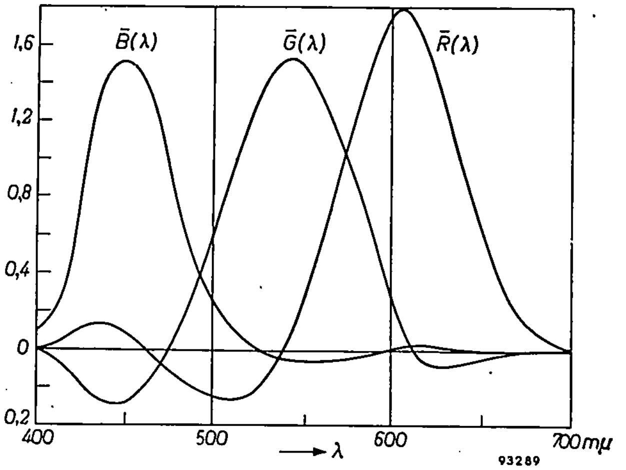

Fig. 9. For exact colour reproduction the quantities $\overline{R}(\lambda)$, $\overline{G}(\lambda)$ and $\overline{B}(\lambda)$ from (12) must, according to the theory, depend upon λ as shown.

In so far, however, as the spectral sensitivity of the camera tube in question, in combination with the selective operation of the dichroic mirror or mirrors, is not entirely in accordance with the curves in fig. 9, an attempt must be made to correct the error by introducing a filter in the path of the light. *Fig. 10* illustrates this for green light. The curves D_b and D_r in fig. 10*a* represent the transmission of the blue and red reflecting mirrors respectively. Multiplying the one ordinate by the other produces the curve D_bD_r, which shows how the light is transmitted which passes through both mirrors. It can be seen that a fair amount of blue is transmitted in addition to the required green. This will no longer happen if the light is also passed through the filter whose transmission characteristic D_f is given in fig. 10*b*. This figure also gives the transmission characteristic $D_bD_rD_f$ of the two mirrors and the filter together. The spectral sensitivity of the camera tube itself (a vidicon) is given by the curve V in fig. 10*c*, the total spectral sensitivity of the tube in combination with the two mirrors and the filter by the curve $D_bD_rD_fV$. The latter is

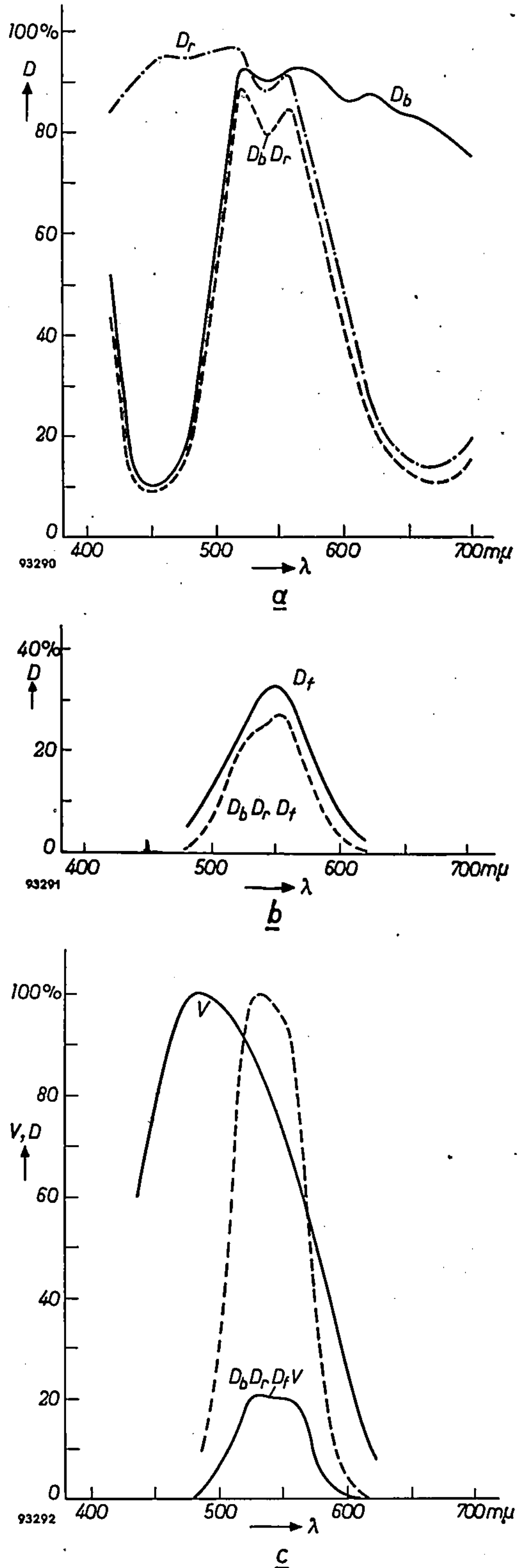

Fig. 10. *a*) Transmission characteristic of dichroic mirrors as a function of λ: D_r for a red reflecting mirror, D_b for a blue reflecting mirror. D_bD_r applies to light transmitted by both mirrors.
b) D_f transmission of a correction filter ("Kodak" 52). $D_bD_rD_f$ represents the transmission of both mirrors and the filter together.
c) V spectral sensitivity of the "green" camera tube (vidicon). $D_bD_rD_fV$ represents the total spectral sensitivity of camera tube with mirrors and filter. This has been re-drawn as a broken curve with its peak at 100%.

also drawn in such a way that the peak coincides with 100%. This curve represents an adequate approximation to the theoretical form of $\overline{G}(\lambda)$ in fig. 9.

Transmission systems

The conveyance of the three video signals in fig. 7 from the transmitter to the receiver by means of a carrier wave, on which they are modulated, requires in all a bandwidth three times as large as that needed for transmitting a monochromatic signal of equal definition. Obviously, ways and means have been sought to reduce this bandwidth without impairing the quality of the picture.

To begin with, we can see what the effect is of reducing the bandwidth of one or more of the primary signals. Slightly reducing the bandwidth of the green results immediately in a noticeable loss of definition. In the case of red, a slight reduction of bandwidth is much less disturbing. The bandwidth of the blue can be appreciably reduced before it adversely affects a normal picture. It appears that the bandwidth can be more limited the less the signal concerned contributes to the luminance ($H = 0.30\ R + 0.59\ G + 0.11\ B$).

An even smaller total bandwidth can be used if, instead of the signals R, G and B themselves, we transmit three mutually independent linear combinations of the signals, one of which is the luminance signal $H = 0.30\ R + 0.59\ G + 0.11\ B$; the two other linear combinations we shall call S_1 and S_2. At the receiving end the signals R, G and B are restored by forming the correct linear combinations of H, S_1 and S_2. The signal H determines the total luminance at the reproduction end, while S_1 and S_2 influence only the colour. The impression of sharpness of the resultant picture is determined primarily by the bandwidth of the luminance signal H. The bandwidth of the colour signals S_1 and S_2 can be strongly reduced without noticeably impairing the quality of the picture.

As already remarked, the luminance signal corresponds entirely to a normal monochromatic video signal. If the Gerber standard (the 625 line system of the Comité Consultatif International des Radiocommunications (C.C.I.R.)) is used with a video bandwidth of 5 Mc/s for the signal H, a bandwidth of, say, 1 Mc/s for S_1 and S_2 will be sufficient for producing a good colour picture. The total bandwidth will then be $5 + 1 + 1 = 7$ Mc/s, as against $5 + 5 + 5 = 15$ Mc/s if R, G and B are to be given the full bandwidth.

The system described, with the signals H, S_1 and S_2, is represented schematically in *fig. 11*. The colour signals at the input of the receiver *Rec* lack the components with frequencies above 1 Mc/s; to distinguish them from the others they are shown in square brackets: $[S_1]$ and $[S_2]$. Because of this difference, the primary colour signals R^*, G^* and B^*, which are formed in the receiver, are not the same as the original primary colour signals R, G and B. The components of R^*, G^* and B^* above 1 Mc/s do not differ from each other since they all originate from the luminance signal H. In America this is called a "mixed-highs" system. It follows from the foregoing that the resultant picture is a three-colour picture only for the components below 1 Mc/s; the components above 1 Mc/s, which provide the fine details, are added in black and white. In most cases the result is not perceptibly different from a picture composed entirely of three colours with all components up to 5 Mc/s.

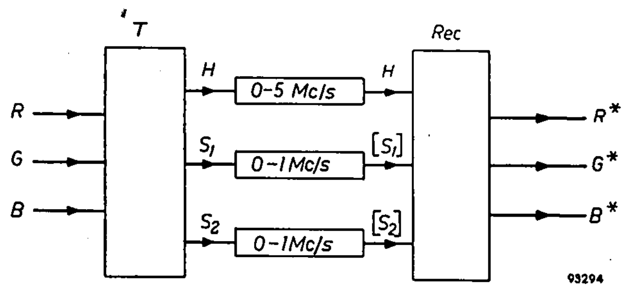

Fig. 11. Colour television system in which the transmitted signals consist of a luminance signal H and two colour signals, S_1 and S_2, which are linear combinations of the signals R, G and B. The signal H is radiated from the transmitter T to the receiver *Rec* with a bandwidth of 5 Mc/s, the signals S_1 and S_2 with a reduced bandwidth, e.g. of the order of 1 Mc/s.

If the following selection is made for S_1 and S_2:

$$\left.\begin{aligned} S_1 &= R - H, \\ S_2 &= B - H, \end{aligned}\right\} \quad . \quad . \quad . \quad . \quad . \quad (13)$$

the signals R^* and B^* are obtained at the receiving end as the sum of the luminance signal H and the colour signal $[S_1]$ or $[S_2]$:

$$R^* = H + [S_1], \quad . \quad . \quad . \quad . \quad (14a)$$

$$B^* = H + [S_2]. \quad . \quad . \quad . \quad . \quad (14b)$$

The signal G^* is produced in a somewhat more intricate manner. It follows from (10) and (13), again indicating the limitation in bandwidth by square brackets, that:

$$0.30\ [R—H] + 0.59\ [G—H] + 0.11[B—H] = 0.$$

Hence

$$[G—H] = -\frac{0.30}{0.59}[S_1] - \frac{0.11}{0.59}[S_2], \quad . \quad (15)$$

and

$$G^* = H + [G - H]. \quad . \quad . \quad . \quad . \quad (16)$$

The four processes to be carried out at the receiving end in order to derive R^*, G^* and B^* from the available signals H, $[S_1]$ and $[S_2]$ are thus:

1) to add H and $[S_1]$: according to (14a), this produces R^*;
2) to add H and $[S_2]$: according to (14b), this produces B^*;
3) to combine certain fractions of $[S_1]$ and $[S_2]$: according to (15), this produces $[G-H]$;
4) to add H and $[G-H]$: according to (16), this produces G^*.

It may be added that (14a), (14b) and (16) can be written as follows:

$$\left.\begin{aligned} R^* &= H + [R] - [H] = [R] + H_{\rm h}, \\ B^* &= H + [B] - [H] = [B] + H_{\rm h}, \\ G^* &= H + [G] - [H] = [G] + H_{\rm h}, \end{aligned}\right\} \quad (17)$$

in which $H_{\rm h} = H - [H]$ represents the luminance signal components with high frequencies (the "mixed highs"). It is evident from (17) that R^*, G^* and B^* contain, as we have just seen, the same high-frequency components ($H_{\rm h}$).

If the luminance signal H is modulated, like an ordinary monochromatic video signal, upon a carrier, and this is transmitted in a television channel together with a sound carrier, it will not differ in any way from a monochromatic transmission. Ordinary television receivers will reproduce such a transmission in normal black and white. This satisfies the requirement of "compatibility", which, in view of the large number of monochromatic receivers in many countries, is of great practical importance.

The question now remaining is, how must $[S_1]$ and $[S_2]$ be transmitted such that a colour television receiver will produce a coloured picture from them? Of course, $[S_1]$ and $[S_2]$ could be separately modulated on carriers outside a television channel. However, in the frequency allotment plan as now generally accepted (for Europe the 1952 Stockholm plan) there is no room for such a procedure. For this reason, means have been sought of accommodating the signals $[S_1]$ and $[S_2]$ in the channels already in use.

System with two subcarriers

The addition to a video signal of a frequency f which is an integral multiple of the line frequency f_l ($f = mf_l$) produces in the television picture a disturbing pattern of vertical strips. The reason for this is that the maxima created in each line by the interfering signal fall exactly below the corresponding maxima in the preceding line (*fig. 12a*), and in all subsequent scans the maxima recur at the same positions. However, if the added signal has a frequency that lies exactly halfway between two multiples of the line frequency: $f = (m + \frac{1}{2})f_l$, there will then be a shift of exactly 180° between the maxima and minima in successive lines (fig. 12*b*).

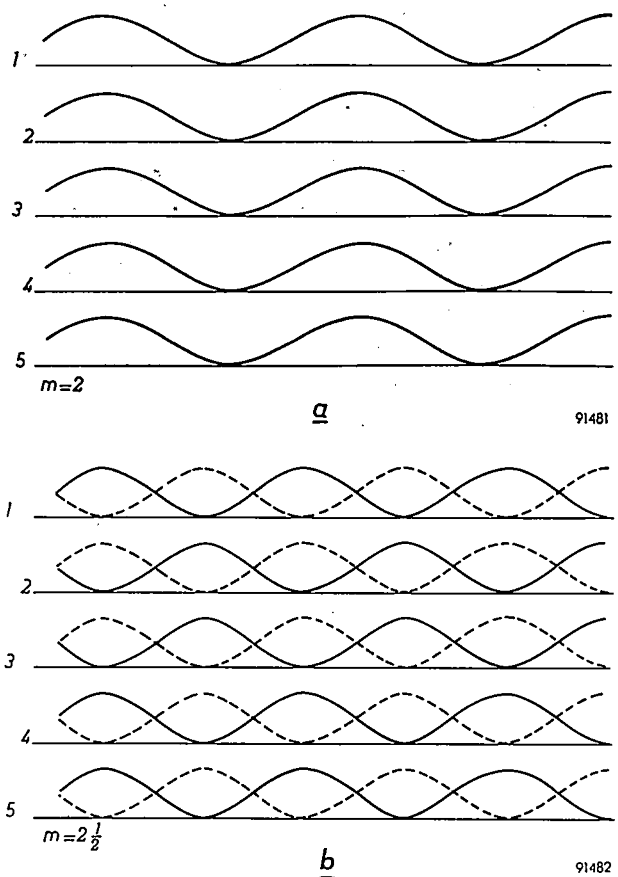

Fig. 12. Variation of luminance (as a result of an interfering signal with frequency f) of the lines 1, 3, 5, . . , 2, 4, . . . of a television scan, *a*) for f = twice the line frequency f_l, *b*) for $f = 2\frac{1}{2}f_l$. In (*a*) the maxima appear directly one above the other and occur at the same place in each successive complete scan. In (*b*) the maxima of the one line appear under the minima of the directly preceding line, and in successive complete scans the maxima and minima occur alternately (full line and broken line).

Moreover, since the total number of lines of a complete picture is uneven, the phase of the added signal is also shifted 180° at each successive scan of the same line (i.e. one picture interval later): at each point where a maximum occurs in the one complete scan, a minimum occurs in the next complete scan. Owing to the inertia of the eye, an added signal of this nature is much less troublesome than a signal with $f = mf_l$. For this reason a signal with $f = (m + \frac{1}{2})f_l$ may permissibly have a fairly large amplitude with hardly any adverse effect on the monochromatic picture, especially if f is not too low.

The above system can be used for adding to the luminance signal a video-frequency carrier, modulated with e.g. the signal $[S_1]$, in such a way as to produce no interference in the monochromatic receiver. For this purpose, this subcarrier must have a frequency of $(m+\frac{1}{2})$ times the line frequency f_l. Although, strictly speaking, the above reasoning no longer holds for moving objects, it has been found in practice that even then the added signal may permissibly have a fairly large amplitude. The second colour signal $[S_2]$ might be modulated on a second subcarrier.

This was the manner of operation of one of the two systems which, in April 1955 and April 1956, were demonstrated in Eindhoven for the C.C.I.R. investigatory committee for colour television [7]).

N.T.S.C. system

Another possibility, which was also demonstrated in Eindhoven on the same occasion, has been developed more especially in America, where it is used for the present colour-television transmissions. In this "N.T.S.C. system" (so called after the National Television System Committee, which coordinates colour television work in the United States [8])) the signals $[S_1]$ and $[S_2]$ are again modulated on a subcarrier, but in this case the two subcarriers have the same frequency and are shifted 90° in phase with respect to each other:

$$[S_1]\cos\omega_a t+[S_2]\sin\omega_a t,$$

in which ω_a is the angular frequency of the subcarriers.

To separate $[S_1]$ and $[S_2]$ in the receiver it is now not possible to use conventional amplitude detectors, since these respond only to the total amplitude $\sqrt{[S_1]^2+[S_2]^2}$. Separation can, however, be effected by means of synchronous detection. If we multiply (e.g. with the aid of a mixer tube) the total colour information signal

$$[S_1]\cos\omega_a t+[S_2]\sin\omega_a t$$

by $2\cos\omega_a t$, we then obtain:

$$\begin{aligned}2[S_1]\cos^2\omega_a t+2[S_2]\sin\omega_a t\cos\omega_a t &= \\ =[S_1]+[S_1]\cos 2\omega_a t+[S_2]\sin 2\omega_a t.\end{aligned}$$

Provided the auxiliary carrier frequency $\omega_a/2\pi$ is high enough (greater than the bandwidth of $[S_1]$ and $[S_2]$), all that remains after passing through a suitable low-pass filter is the signal $[S_1]$. The signal $[S_2]$ is restored in an analogous way.

For this synchronous detection the auxiliary signals $\cos\omega_a t$ and $\sin\omega_a t$ are needed in the receiver. These are obtained from a wave train ("burst") sent out by the transmitter after each line synchronization pulse; the burst consists of a specific number (e.g. eight) cycles of the subcarrier frequency and has a specific phase, e.g. the cosine phase. In the receiver the burst can be used to synchronize an oscillator, from which the one auxiliary signal is obtained and, after a 90° phase shift, the other.

If, as under (13), S_1 is taken as $R-H$ and S_2 as $B-H$, each possibly multiplied by a constant factor

$$S_1=\alpha(R-H)\ \text{and}\ S_2=\beta(B-H),$$

then the total colour information signal (see *fig. 13*) is:

$$S=\alpha(R-H)\cos\omega_a t+\beta(B-H)\sin\omega_a t. \quad (18)$$

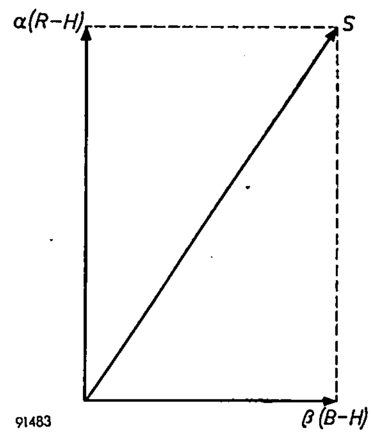

Fig. 13. The total colour information signal S according to the N.T.S.C. system (see eq. 18).

For white (and grey) this signal is zero. An increasing amplitude of the signal, the luminance signal remaining constant, corresponds to increasing colour saturation. A change of phase constitutes a change of the dominant wavelength of the colour.

In the American colour-television transmissions a small refinement is at present being employed, as described below.

The video spectrum according to (18) appears as drawn in *fig. 14a*. f_a is the frequency of the subcarrier wave on which $[S_1]$ and $[S_2]$ are modulated according to the normal double-sideband system. The latter is necessary in order to be able to separate the two colour signals by means of synchronous detection; with single-sideband modulation there would be complete cross-talk between these signals. The largest bandwidth that can be chosen for $[S_1]$ and $[S_2]$ in this system is thus the total video bandwidth (5 Mc/s in the Gerber system) minus the subcarrier frequency f_a.

[7]) J. Haantjes and K. Teer, Compatible colour-television, I. Two sub-carrier system, II. Comparison of two sub-carrier and N.T.S.C. systems, Wireless Engr. **33**, 3-9 and 39-46, 1956 (Nos. 1 and 2).

[8]) D. G. Fink, Color television standards; selected papers and records of the N.T.S.C., McGraw-Hill, New York 1955. Hazeltine Labs. Staff, Principles of color television, John Wiley & Sons, New York 1956.

The refinement referred to consists in giving one of the colour signals a larger bandwidth, by modulating the high-frequency components of the signal with only one sideband upon the subcarrier. At demodulation there is admittedly cross-talk between these components and the other colour signal, but this interference can be removed by means of a low-pass filter.

This system has been adopted in the present-day N.T.S.C. system. For the colour signal with enlarged bandwidth, one might choose between $\alpha(R-H)$ and $\beta(B-H)$, but it is possible to make a better choice. In the expression

$$S = S_1 \cos \omega_a t + S_2 \sin \omega_a t$$

the right-hand term is identical with:

$$(S_1 \cos \varphi - S_2 \sin \varphi) \cos (\omega_a t + \varphi) + \\ + (S_1 \sin \varphi + S_2 \cos \varphi) \sin (\omega_a t + \varphi)$$

for each value of the phase angle φ. The signal S can therefore also be formed by modulating the subcarrier wave in the phase $\cos (\omega_a t+\varphi)$ with the signal $(S_1 \cos\varphi - S_2 \sin\varphi)$ and in the phase $\sin (\omega_a t + \varphi)$ with the signal $(S_1 \sin \varphi + S_2 \cos \varphi)$, while still

$$S_1 = \alpha(R-H) \text{ and } S_2 = \beta(B-H).$$

Experiments have been made to ascertain for what value of φ an enlarged bandwidth of the signal to be modulated on $\cos (\omega_a t + \varphi)$ contributes most to improving the colour picture. With the values of α and β chosen in America the greatest improvement was obtained for $\varphi = 33°$. With this value the colour information signal is:

$$I \cos (\omega_a t + 33°) + Q \sin (\omega_a t + 33°),$$

in which

$$I = \alpha(R-H) \cos 33° - \beta(B-H) \sin 33° \quad . \quad . \quad (19)$$

thus gets a somewhat larger bandwidth than

$$Q = \alpha(R-H) \sin 33° + \beta(B-H) \cos 33°. \quad . \quad . \quad (20)$$

The upper sideband is moreover limited to the bandwidth of the signal Q. The result is illustrated in fig. 14*b*.

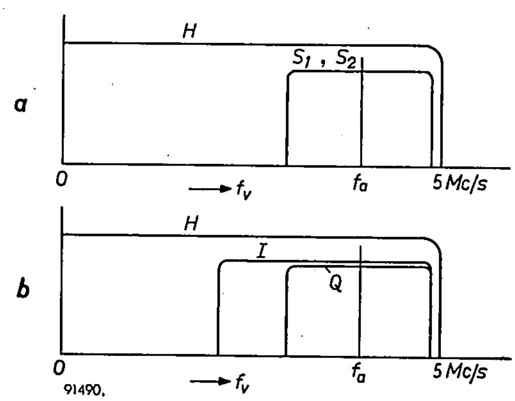

Fig. 14. *a*) Video-frequency spectrum in which the two colour signals, S_1 and S_2, have equal bandwidths. Each is modulated with two sidebands of limited width on a subcarrier with video frequency f_a; the two subcarriers are 90° out of phase with each other.
b) Video-frequency spectrum in which the two colour signals, I and Q, have different bandwidths (see (19) and (20)). In this case, I can have a wider lower sideband. This system is employed in America.

Variation in the amplitude of the signal I results in a change of colour along the line which in the colour triangle connects orange with blue-green (the line I in *fig. 15*). Information on colour variation along this line appears to be more important than information on variation along the green-purple line, which relates to variation of the signal Q (the line Q in fig. 15).

Gamma correction

As remarked, the relation between the luminous flux Φ and the control voltage V (measured from the cut-off point) is not linear in conventional picture tubes. By approximation this relation is a power function:

$$\Phi = c\, V^{\gamma}, \quad . \quad . \quad . \quad . \quad . \quad . \quad (21)$$

in which the exponent γ lies between 2 and 2.5. It is necessary to make a correction for this non-linearity (gamma correction), as otherwise impermissible errors would arise in the reproduction.

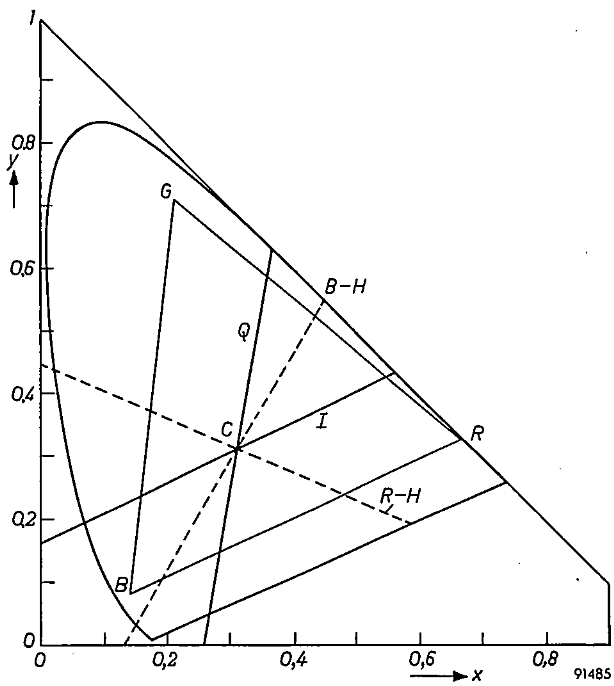

Fig. 15. Variation of signal I (eq. 19) causes a change of colour along line I in the colour triangle; variation of signal Q (eq. 20) causes a change along line Q. The colour points of the primary colours are R, G and B. The signals S_1 and S_2 correspond to the lines R—H and B—H.

If every receiver were to be provided with such a correction, the price of the receivers would be considerably increased. For that reason, the gamma correction is made in the transmitter, by including a non-linear element (gamma corrector) in each of the three channels (red, green and blue) before the signals are made up into luminance and colour signals. If the expression (21) applies at the reproduction end for each of the three primary colours, the relation between the output voltage

V_o and the input voltage V_i of the gamma corrector must be:

$$V_o = c' \, V_i^{1/\gamma},$$

if, at least, the camera tubes supply a signal that is linearly dependent upon the luminance L_i on the photo-cathode. If this is not the case, the gamma corrector is made such that the relation between V_o and L_i is:

$$V_o = c'' \, L_i^{1/\gamma}.$$

If this is done in each colour channel, we obtain, instead of the signals R, G and B, the corrected signals

$$R' = R^{1/\gamma}, \quad G' = G^{1/\gamma}, \quad B' = B^{1/\gamma}.$$

The new luminance signal H' now becomes (eq. 10):

$$H' = 0.30\, R' + 0.59\, G' + 0.11\, B',$$

and the new colour signal S' (in the N.T.S.C. system) becomes

$$S' = \alpha(R' - H') \cos \omega_a t + \beta(B' - H') \sin \omega_a t.$$

In the receiver these signals are treated exactly as described above for the uncorrected signals, so that R', G' and B' are the signals which are fed to the reproduction device.

Gamma correction at the transmitting end has, it is true, certain disadvantages, which we shall not go into here; on the other hand it is indispensable for ensuring good reproduction without making the receivers all too complicated.

Summary. This introductory article on colour television begins by recapitulating some colorimetric concepts, such as additive colour mixing and the chromaticity diagram (colour triangle). The principle of colour television is then explained with reference to a system containing three camera tubes and three picture tubes, considered at this stage with separate transmission channels for the three primary colours (red, green and blue). The splitting of the incident light into the three primary colours at the transmitting end, and the combination of the three (projected) pictures at the receiving end, can be effected by means of dichroic mirrors.

As regards transmission the main problem is to limit the bandwidth. A discussion follows of a system with two subcarriers and of the N.T.S.C. system used in America. Finally, the author deals with the gamma correction needed in view of the non-linear relation between luminous flux and control voltage in picture tubes.