CHAPTER 1

Introduction

Some of the earliest thoughts of television included the concept of pictures in full color. Inventions were made by the score and were almost universally confined to paper designs (*1*). The few demonstrations of the 1920's used scanning-disc techniques and showed only that the concepts of color, in terms of tristimulus phenomena, were applicable to electrical control. By 1940, however, black-and-white electronic television with camera tubes and a cathode-ray-tube reproducer was well advanced.

That year, an impressive demonstration was made (*2*) of color television using a synchronized rotating color wheel in front of a black-and-white picture tube. The picture was then modulated sequentially in accordance with the three color-component luminances. Unfortunately, mechanical limitations of size and rotational speed are sufficiently severe so that this type of field-sequential color picture reproduction has disappeared. (Rotating color filters, however, remain useful in small sizes, for use with TV cameras, such as those used in early television from the moon.)

The major challenge that faced inventors and promoters of color television was to find a picture reproduction method that compared in simplicity to that of the black-and-white cathode-ray tube.* Optical superposition of pictures from three tubes was tried and is still in use in projection systems. For direct viewing on a single picture tube, however, the first successful demonstration was made in early 1950 using an internal structure that is commonly known as the shadow mask (*3, 4*). This type of tube employs three electron beams which pass through a common deflec-

* For a bibliography and description of important early work not otherwise referred to herein, see Herold (*1*).

tion yoke (*5*), with one beam for each primary phosphor color, i.e., red, green, and blue. The beams are "shadowed" by a perforated metal mask so that each beam can strike but one color of phosphor. The shadow-mask tube was intensively developed in the 1950–1957 period and many basic improvements were made. Of these, the most important were the use of a curved mask (*6*) (initially, it was flat), the use of photodeposited phosphor dots (*7*), better phosphors, and refinements of technology in manufacture that led to better uniformity and lower cost. However, two different single-beam tubes were also developed. In the one, then known as the Lawrence tube (*8*), vertical phosphor stripes were used to form a line screen that was spaced from a wire grill on which a high-frequency potential was impressed to deflect the beam from one phosphor color element to another. The other single-beam type also used vertical phosphor stripes and a single beam, but required no mask or grill (*9*). Instead, the phosphor screen supplied an index signal to switch the beam control to whichever color signal corresponded to the phosphor being struck. This early period also led to proposed variations of the three-beam shadow-mask tube, some using wire grills and line screens instead of perforated masks (*10–12*). Although the principles of other types of color tubes still survive in the laboratory, the three-beam shadow-mask method outdistanced its competitors and achieved commercial success (*13, 14*).

From 1955 to 1967, about 15 million shadow-mask color picture tubes were made, and manufacture of a highly satisfactory product became routine. Round metal envelopes were replaced by glass ones (*15*); these were then changed to a rectangular shape, picture size was increased, and deflection angles (total swing across diagonal axis) were changed from 70° to 90° thereby improving picture sharpness and decreasing tube length (*16–18*). By 1973, color television was well established in most industrially sophisticated countries of the world, and of the order of 20 million color picture tubes were being made each year, worldwide. Improvements and innovations were added to the basic shadow-mask principle. Among the more important were a trend to even larger deflection angles of 110°, and use of a black matrix that surrounded each phosphor element so as to reduce reflected light and increase contrast (*19, 20*). Coupled with new, highly efficient, color phosphors, some using rare-earth elements (*21, 22*), these improvements permitted major increases in picture brightness. For the electron beam, there were modifications in the gun for both the most common triangular arrangement and also for the in-line arrangement of beams (*23, 24*). Although the most common aperture shape in the shadow mask remained circular, several commercial types were introduced that used vertically elongated apertures (*24, 25*). The phosphor screens in these types resembled the vertical line screen of the early single-beam tubes mentioned above, although the principles were those of the shadow mask.

In this discussion, we shall first review the fundamental principles that underlie any color picture tube and classify the several methods that make such a tube possible. Because of its predominance, greatest attention will thereafter be given to the shadow-mask method and its principles of design and manufacture. Some of the other methods, particularly those that do not appear to have been entirely abandoned, will also be described, but in less detail. The emphasis, throughout, will be on tubes for full color display, i.e., tubes providing the three primary colors and a luminance range (gray scale) sufficient for good contrast and satisfactory picture rendition. It is evident that there are applications that are considerably less demanding; for example, two-color systems or systems for alphanumerics or line drawings in which color is used but no "gray scale" requirement is imposed. Such displays can be made with the cathode-ray tube methods herein covered, but permit employment of many other methods not included in the present volume.

CHAPTER 2

Requirements of a Color Picture Reproducer

The variables in a reproduced picture may be briefly listed as size and shape, sharpness and resolution, brightness (luminance), contrast, and color hue and saturation. Some of these are related in that one can be traded off for another in the design of a reproducer. Also of importance are such matters as white balance and uniformity, the "gamma" or degree of nonlinearity of light output vs. electrical input, and spatial coincidence when separate color images are superimposed. Each of these factors will be discussed in turn.

The size of a reproduced color picture is one of the most flexible of the requirements. Direct-view picture tubes, as of 1973, range in size from 266 mm (9V) diagonal to as much as 667 mm (25 V). Projected pictures can be much larger. Because of the well-known psychological principle of "size constancy" with distance (*26*), large pictures are more pleasing than small ones, even when the latter are viewed at shorter distances and subtend the same angle at the eye. Since present television systems transmit a rectangular picture with a 4:3 aspect ratio, the reproduced picture should conform to this aspect ratio to avoid picture distortion or loss of information. It is common practice to overscan or use a raster that exceeds the phosphor screen size which, especially for circuits employing tubes, may be needed to provide for possible raster shrinkage with use of the receiver. Overscanning, of course, causes a loss of picture information and brightness, but makes objects appear slightly larger.

Sharpness and resolution of a color picture reproducer are best discussed in terms of the modulation transfer function (MTF), i.e., the sine-wave luminance response curve plotted against spatial frequency (*27, 28*).

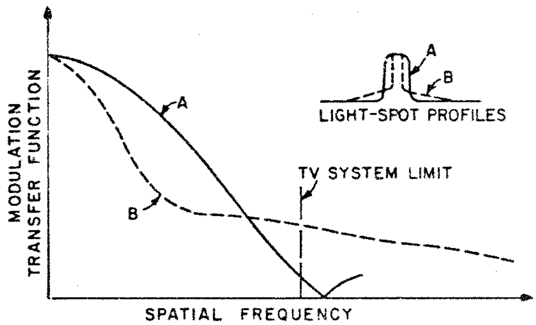

Fig. 2.1. Response curves for two hypothetical picture tubes. Tube A will have a much sharper picture even though it will not show scanning lines as well and will have only half the resolution of Tube B.

Sharpness is the subjective effect on the viewer of step-function changes in scene luminance, and resolution is the ability to distinguish detail. In terms of the MTF, limiting resolution is determined by the highest spatial frequency that is discernible, and sharpness is determined by the shape of the MTF response at frequencies below the resolution limit. In television, there is a resolution limit imposed by the system itself. In addition, the eye has its own MTF and resolution limit that depends on the viewing distance and the pupil diameter. A picture reproducer need not have resolution beyond the limit of the television system and the eye. Below this limit, however, the reproducer should have a luminance MTF as close as possible to 100%. For most existing color TV systems, the upper limit for the resolution of the system is between 350 and 600 TV line numbers (equivalent to 175 to 300 line pairs per picture height). The smaller numbers are in the horizontal direction and the larger are the extreme imposed by the scanning lines, neglecting the "Kell" factor reduction in vertical resolution (*29*) by 0.7. With bright pictures and close viewing distance, the luminance MTF of the eye is not limiting and can be disregarded. Because the eye is relatively insensitive to color in small detail, the MTF for color is of even less importance. However, known reproducers tend to have equal MTF responses for the three primary colors, i.e., they exceed the perceptual requirements substantially.

To illustrate the significance of the above, Fig. 2.1 shows modulation transfer functions for two hypothetical color TV picture tubes. Tube A has a nearly ideal flat-topped light-spot profile, while Tube B has a much smaller central spot but wide flaring skirts similar to a halo. The light-spot profile and the MTF response curves are Fourier transforms of each other. Although reproducer B has much higher limiting resolution than A, and

A has poorer responses at the scanning line frequency, A will nevertheless have a much sharper picture and would be preferred over B for the TV system limit shown in Fig. 2.1, and possibly for any other system as well (*30*).

A projected still color picture continues to give greater satisfaction as highlight brightness is increased up to some few thousands of candela per square meter.* In television systems employing 50 fields/sec, flicker begins to be annoying as brightness is increased above approximately 200 cd/m² when using the exponentially decaying phosphors commonly used in direct-view picture tubes. Somewhat brighter pictures would be tolerable if the "on" time for a picture element were more nearly square-topped. With 60-field/sec systems using phosphors, flicker becomes annoying with peripheral vision at something over 350 cd/m² but can be tolerated with on-center viewing up to considerably more than that (*31*). Such high luminance values are readily obtained on direct-view tubes but are much more difficult to achieve in projection systems. The eye is so adaptive, however, that pictures of only a few candela per square meter are pleasing if viewed in sufficient darkness. On the other hand, in home television the ambient illumination is often appreciable and high brightness is decidedly advantageous. Under such conditions, brightness and contrast are closely related, because the ambient light is often reflected and diffused from the face of the picture tube and raises the black level.

Contrast is one of the most perceptible factors in a reproduced picture, particularly when the picture is in color. If the ratio of the highest luminance to the lowest is reduced because of ambient light, color saturation also is reduced and the picture appears "washed out" in both color and in contrast (*22*). To improve contrast under ambient light, brightness has often been sacrificed, as with direct-view picture tubes in which a neutral-density absorbing glass is used to reduce the effect of the ambient illumination. Such glass is effective because the ambient light is diffused by the white phosphor and passes through the faceplate twice, while light emitted from the phosphor passes through only once. In more recent designs of shadow-mask tubes, a black matrix surrounds the phosphor dots (*20*), thereby reducing the need for absorption in the glass and permitting greater brightness. In darkness, the contrast ratio of direct-view picture tubes is limited by electron and optical scattering to the order of 50:1, but this contrast ratio can give an excellent picture. Projected pictures usually have a lower contrast ratio than this. Many home direct-view picture tubes are viewed with about 250 lux ambient illumination and have a contrast ratio of under 10:1; yet this condition is accepted.

* 1 footlambert (fL) = 3.43 cd/m² = 3.43 nits.

Human perception of absolute color hue and saturation is extremely diverse, both because of individual variations and because the eye is primarily conscious of color differences and not of absolute color. On the other hand, a color television system must be designed to permit a relatively high degree of color fidelity. Fortunately, only three primary colors are sufficient; if there is a match between the primaries at the camera end and those used for the picture reproduction, accurate colors will result and perceived color differences will be equally correct. Using the International Commission of Illumination chromaticity diagram shown in Fig. 2.2 (known as the CIE diagram), the entire range of colors observed by the normal eye is found within the horseshoe-shaped figure (*32*). Any color may be represented by the x, y coordinates in this diagram, with the hue angularly displaced around the central "white" region and the saturation increasing from the center out to the curved border. Combinations of two colors lie on the line joining their two coordinate points. The three primary points specified for most color television systems throughout the world are indicated on the diagram by open circles and, except for a region in the green-cyan range, cover enough of the area to produce very satisfactory color accuracy. The "white" point for the system is also indicated and corresponds closely to the radiated color of a blackbody at 6500°K, an approximation to daylight.

For the reproducer to do maximum justice to the system, its primary colors should correspond with the system primaries. Phosphors available for color picture tubes permit this, but not always under highest efficiency, i.e., maximum brightness. For this reason, compromises are usually made, particularly for the red and the green phosphors. The adaptability of the viewer and the propensity of most viewers to trade off some color fidelity for brightness have led to acceptance of picture-tube colors displaced from the specified system, as indicated in Fig. 2.2 by the crosses and the dashed triangle joining them. The greatest loss of saturated color is in the cyan-to-green region, but a brighter picture is obtained with excellent color fidelity in the red-to-yellow-green region. Fortunately, it has been shown that few objects and scenes require the missing region of color (*22*). In the United States, because there has been a consumer preference for a very bluish "white" for monochrome television, color receivers are normally adjusted for a white point close to the 9300°K blackbody point, another expedient deviation from the system. It is likely that, as color becomes predominant in television and improvements are made in phosphors, the white point and the three primary colors will eventually match those for which the system was designed.

With all systems of color television, the reproducer with its associated circuits must be controllable independently for each color primary. The

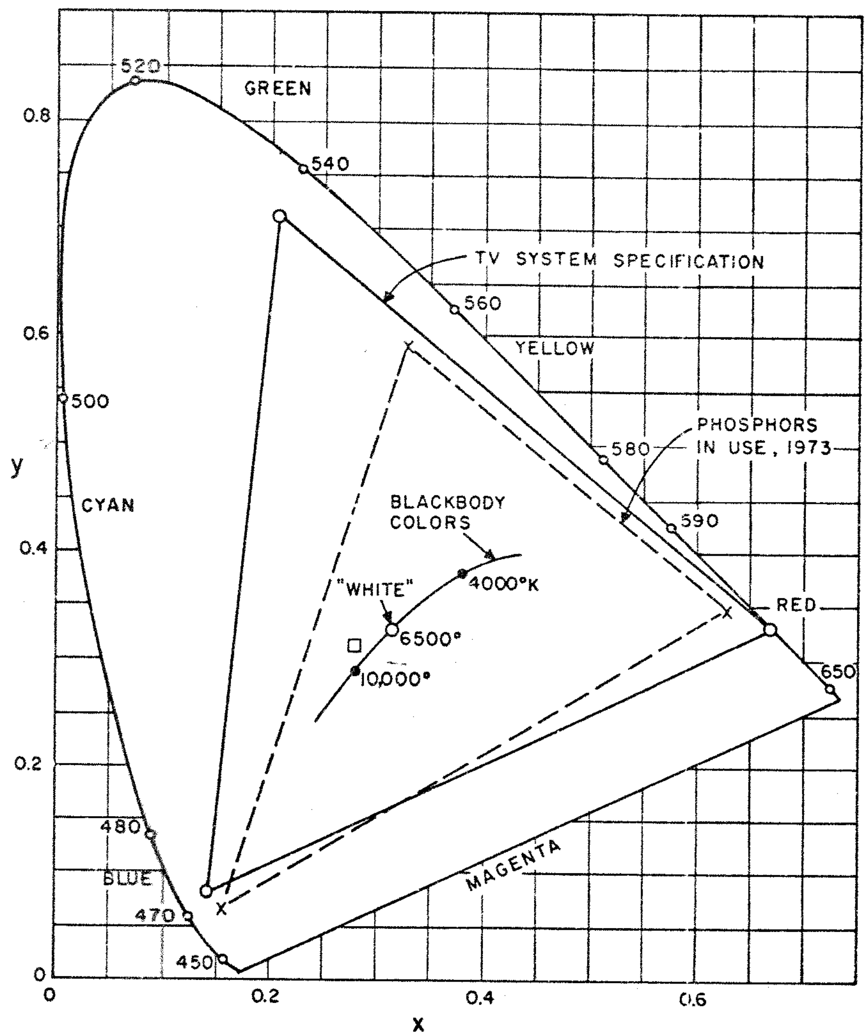

Fig. 2.2. The CIE color diagram. The specified color system primaries and white point are shown by the open circles. Typical United States color picture tube phosphors are shown by crosses and the frequently used white by the square.

color produced with any one or any fixed arbitrary combination of the three primaries must be uniform over the entire picture area. The white balance, i.e., the combination of the three primaries to produce white, must be either controllable or fixed at a point in the CIE diagram suitable for the system. In addition to these requirements, the white balance and the color uniformity must be invariant with the light output, from dark to bright. For example, in a picture tube that uses three electron beams, one for each color, this requirement imposes close similarity in the control characteristic for each of the beams. It is much more easily achieved if the phosphors have spectral characteristics and efficiencies such that the white balance is obtained with equal current from each beam.

All commonly used color television systems, in order to improve signal-to-noise ratio (and for other reasons as well), transmit an electrical signal that is not linearly proportional to the desired light output at the reproducer. Instead, the signal has an amplitude compression that corresponds to a fractional power law, where the exponent is approximately 0.46, otherwise known as a "gamma" of 0.46. In a color television system, true color fidelity requires restoration of proportionality so that the reproduced luminance is proportional to the original luminance. For this reason, the transmitted signal, with its gamma value of 0.46, must be expanded by a 2.2 power law (the reciprocal of 0.46) to produce the brightness variations in the reproduced picture. Fortunately, this expansion is easily done by the electron gun used for most color television picture tubes. Experimentally, it is found that practical gun designs have a nonlinear light output vs. control signal characteristic that approximates a power law with an exponent somewhere between 2 and 3. Because the contrast range in practical systems is limited to about 50:1, this is transmitted in compressed form by a voltage range of $50^{0.46}:1$ (or about 6:1). Thus, the signal range over which the picture tube gun must approximate the desired 2.2 power law is not large. Furthermore, the signal amplitude is adjustable (this is done by the contrast control and video drive circuits in TV receivers), and so is the background level (adjustable by the brightness control). When these degrees of freedom are included, it is found that picture-tube guns can be made to approximate the desired gamma value of 2.2 within the signal limits of importance. Practical guns used in picture tubes, therefore, are capable of faithful color rendition, even though they may not have constant-power-law characteristics. In any picture reproducer having a radically different relation between light output and signal than that of the cathode-ray tube, electrical gamma compensation must be introduced.

The last requirement of a color picture reproducer is pertinent only when the viewed picture is a superposition of three separately generated pictures as, for example, in the common shadow-mask tube, or in three superposed projected pictures. This final requirement is coincidence. The separate images should be in coincidence, preferably up to the limit of resolution, and they must stay in coincidence at any brightness level and over the entire picture area. Fortunately, minor deviations from coincidence are not noticeable at larger viewing distances. The effect is also less critical than one might suppose because so much of the luminance information is contained in the green signal, and the color acuity of the eye is less for small detail in the other colors. In practice, coincidence has not been a problem when color television pictures are reproduced by well-designed and well-adjusted shadow-mask tubes or separately projected color images.

CHAPTER 3

Classification of Methods for Color Picture Reproducers

An earlier paper (*1*) reviewed many color picture tube proposals in the patent literature prior to 1951. The ideas are extensive and most of them are no longer considered feasible. In this chapter, we shall confine ourselves to a general classification of methods that are still of interest, putting them in perspective. Detailed description follows in Chapters 5–9 and the reader is also referred to these chapters for literature references.

The most obvious way of producing a color television picture is to superpose optically three separate pictures, one for each primary color. This method is still used in large-screen projection systems, and it was once even used for experimental home receivers. However, a large and unattractive bulk is needed for mirrors or for other optics. Much more interesting are methods in which the superposition of the three pictures is direct, as on the face of a picture tube. One way to accomplish this is to segment color phosphors into small triplet groups or triads, with one phosphor element of each triad for each primary color. When such a screen is viewed at a distance, the red, green, and blue pictures appear superposed even though they are actually produced by closely spaced fine lines or dots. A second class of device uses a nonsegmented or continuous phosphor screen that has several color layers. The same spot on the screen can produce any primary color (or a combination of colors) by varying the electron beam velocity, hence its penetration into the layers.

3.1 Segmented Phosphor Screens

Returning to the first class of device with its segmented phosphor screen, it is necessary to have independent control of each color. If scanning of the electron beam were sufficiently accurate, one could assume that the beam position would be exactly determined at any given time. In that case, one could rely on the scanning linearity to register the electron beam with the proper color. The incoming color signals could then be programmed so that the beam intensity is exactly correct for a given primary at exactly the time the beam is striking that color phosphor. Attempts were made to do this, both with horizontal stripes of alternate red, green, and blue (R, G, B) phosphor and with similar vertical stripes, but it was much too difficult to get the required accuracy. For this reason, all the surviving methods that employ a segmented screen make use of some auxiliary structure at or near the phosphor screen and in accurate register with the color triplets, whether they be stripes, dots, or any other shape. As we shall see, this structure eliminates any need for extremely accurate scanning.

3.1.1 The Shadow-Mask Tube

The auxiliary structure at the screen that has been most successful is that of the shadow-mask tube. In this type of color picture tube, three closely spaced electron guns produce beams that pass through holes or slits in a metal mask where the mask is spaced apart from but close to the phosphor screen. Mask and phosphor screen are at the same voltage, so the electrons travel in essentially straight lines. At the point where the electrons from one of the guns impinge on the screen, one of the color phosphors is deposited in a spot or line that approximates the size of the mask opening. All other parts of the phosphor screen are in the "shadow" of the phosphor mask, as far as this one gun is concerned. The guns are separated in such a manner that the three color phosphor segments do not overlap. Each gun controls the light from only one phosphor, so the tube can be thought of as three cathode-ray tubes, physically interlaced in the same envelope and with the same deflecting system. The tube is useful for any incoming television signal that can be separated into red, green, and blue components, one for each gun, and so has wide flexibility.

The principle of the shadow-maks tube is illustrated in Fig. 3.1 in two different forms. In one case, round openings in the mask are used and the phosphor dots are also round. This structure is the original form of the shadow mask and as of 1973 is still predominant. Another form uses a mask having elongated slits and a faceplate on which vertical lines of alternately red, green, and blue emitting phosphors are deposited. Although only center lines are shown in Fig. 3.1, in each case there are three electron beams that originate from three independent electron guns (one for each color). These guns are closely spaced, and the electron beams make an angle of only 1° with each other as they pass through the shadow mask. A single system deflects all three beams, and the three beams converge to approximately the same point on the screen. With the circular hole and its round phosphor dot, the most common gun configuration is a triangle or delta arrangement. The slit shadow mask uses in-line guns. Since the electron-beam spot usually covers more than one opening in the shadow mask, the phosphor dots or the phosphor line widths are smaller than the electron beam spot. As a result, screen structure is not generally noticed in normal picture viewing.

In both the hole and the slit arrangements, it is necessary that each electron beam strike only one color at all deflection angles. The tolerance built in to assure this condition requires either that the phosphor dot or line be slightly larger than the mask opening (known as positive tolerance), or that each individual phosphor area be smaller than its corresponding mask opening (known as negative tolerance). In either case, a slight shift

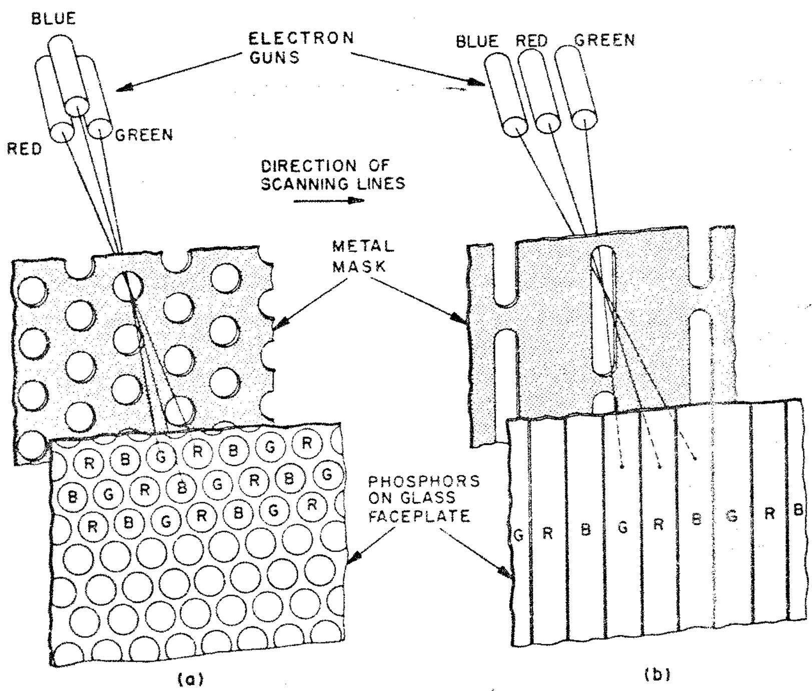

Fig. 3.1. Schematic of two shadow-mask systems: (a) with round holes and delta guns and (b) with slit openings and in-line guns. In actual tubes, mask and faceplate are curved.

in beam landing does not change the brightness or the color purity. To improve contrast, all screen areas not covered by phosphor may be coated with black inert material (called a matrix); alternatively, the faceplate glass may be tinted with a neutral gray. The brightest pictures are obtained with nearly transparent glass using a black matrix. Other factors that contribute to brightness are a high voltage on the phosphor screen and a large deflection angle. The large angle allows a reduction in gun-to-screen distance which minimizes spot size at high currents.

The complexities of process technology and the details of shadow-mask tube design defy simple description and the reader is referred to Chapter 5 in which these matters are discussed in detail. Suffice it to say here that, because the electron shadows and the shadows of a light beam are almost identical, the phosphor areas can be located by photodeposition, and this has been the clue to successful mass production.

3.1.2 The Focus-Mask Tube

It is evident from Fig. 3.1 that much of the electron beam is intercepted by the shadow mask and never reaches the phosphor screen to produce light. If the phosphor-dot screen shown in Fig. 3.1 is operated at a higher voltage than that of the shadow mask, the electrons will no longer travel in straight lines beyond the mask and will be focused into a spot smaller than the shadow-mask opening. Each opening acts as a focusing lens. Because of this focusing action, the round apertures in the mask of Fig. 3.1 can be enlarged and yet the electrons will not exceed their intended phosphor area. The larger apertures greatly reduce the beam current intercepted by the shadow mask and increase picture brightness correspondingly. In some forms, the focus mask consists of a grill of vertical wires (with corresponding vertical phosphor stripes) rather than a mask with round apertures (*10, 11*). This variation is sometimes known as the focus-grill tube.

Unfortunately, the simple focus-mask principle has disadvantages and has not had wide use as of 1973. Because the mask apertures are much larger than the phosphor elements, simple photodeposition techniques for forming the phosphor screen are no longer applicable, and mass production with uniform performance is much more difficult. Another disadvantage is the undesired excitation of phosphor areas by secondary electrons scattered from the mask (or generated at the phosphor screen and reflected back to other areas). These undesirable effects are caused by the phosphor screen being operated at a much higher voltage than the mask. In a simple focus-mask tube the scattered electrons reduce the contrast and dilute the colors; but with more complex structures that use

CHAPTER 4

Limiting Factors on Screen Brightness in Picture Tubes

The preceding chapter described a variety of ways in which color television pictures may be reproduced. Most of them use cathode-ray-tube technology based on deflection of an electron beam that strikes a phosphor screen. These share some general limitations. There are other limitations that depend on the type of tube.

To produce a sharp and bright picture, an electron beam should carry large current and form a small spot at the phosphor screen, resulting in high current density. The current must also be controllable down to a negligible value by a signal voltage. An electron gun producing such a spot generates a small source of electrons that is imaged by an electron lens into the spot on the screen. Since a deflection system of minimum power is desirable, the active deflection volume should be small, requiring a gun of small diameter. If the tube is to be short, the gun cannot be long, and the deflection angle must be large. Some of these requirements are mutually contradictory. This chapter describes the limits to which an optimal compromise is subject. There are some simple fundamentals that lie behind these limits.

(1) The spot at the screen has a current density that depends on that of the cathode, so that maximum cathode current density is desirable.

(2) The spot at the screen becomes smaller and the current density greater when the image distance of the main lens is decreased, corresponding to a larger deflection angle for fixed picture size. A large deflection angle, however, requires large deflection power.

(3) The spot at the screen becomes smaller and the current density greater when the object distance of the main lens is increased or the gun is

made longer. A longer gun increases the tube length and the required electron lens diameter.

(4) The spot at the screen can be made smaller if the lens aberrations are reduced. This requires a large lens diameter and correspondingly large deflection power.

(5) High light outputs and small spot diameters are achieved with high screen voltages; these require large deflection power.

Compromises are obviously needed in any practical design. Additional limitations, calling for further compromises, occur (a) when three beams are used in a single envelope, as with the shadow-mask tube, and (b) when a single gun must reproduce the three primary colors in time sequence, as with the line-screen index tube. The basic limitations on screen brightness for all these cases are the subject of Chapter 4.

4.1 Beam Current Limits in Cathode-Ray Tubes

Functionally, the electron gun in a cathode-ray tube may be regarded as composed of a beam-forming system, which concentrates the electron emission from the cathode so as to form a narrow beam constriction or crossover, and a "final lens," which images the crossover onto the screen, forming the scanning spot (Fig. 4.1). The crossover radius is given by the product of the focal length of the beam-forming region and the ratio of the average initial velocity of the electrons (determined by the effective cathode temperature) to the electron velocity at the crossover. To a first approximation, it is independent of the diameter of the emitting area of the cathode and of the convergence angle at the crossover, which is proportional to this diameter. Similarly, the magnification of the crossover by the final lens is given simply by the product of the ratio of the image and object distances of the final lens and the ratio of the electron velocities at the crossover and at the screen and is thus, also, independent of the convergence angle at the crossover and the cathode area contributing to the beam. It might be concluded that the contributing cathode area and the covergence angle at the crossover could be increased indefinitely without increasing the spot size on the screen. However, the spot diameter is not simply the product of the crossover diameter by the final-lens magnification, but is augmented, as a result of the aberration of the final lens, by an amount proportional to the third power of the convergence angle at the crossover.

As a result, referring once again to Fig. 4.1, there is an optimum convergence angle α_1 for any prescribed spot diameter d_s [Eq. (4.19)]* and the maxi-

* Reference is made to equations appearing at a later point in the text to give the reader insight into the quantitative relationships without demanding familiarity with their derivation.

mum beam current within the spot is proportional to the 8/3 power of the spot diameter [Eq. (4.21)].

The maximum spot current for prescribed spot diameter also depends on the aberration constant of the final lens. This aberration constant Cf_2 is primarily a function of the clear or unobstructed diameter D of the electrodes (or pole pieces) of the final lens and of its focal length f_2 and depends relatively little on the lens type. Thus, substitution of the relationship between the aberration constant and the clear diameter and focal length for a favorable and commonly used lens type (the equidiameter-cylinder accelerating lens) leads to a generally valid expression for the upper limit to the beam current set by thermal spread and lens aberrations, in terms of cathode emission j_0, cathode temperature T, screen voltage V, and geometrical parameters [Eqs. (4.22) and (4.24)].

Thermal spread and aberration are not the only factors which limit the current that can be concentrated in a spot of given dimensions. Another factor is the mutual repulsion of the beam electrons, or the space-charge forces, in the drift region between the gun and the screen. The current limit established by space charge becomes significant at relatively high beam current values and is proportional to the 3/2 power of the screen voltage, the square of the convergence angle in the drift region, and to a universal function of the ratio of spot diameter to beam diameter at the entrance to the drift region [Eq. (4.27)].

The results of the mathematical analysis which follows are given not merely in the form of equations but are also represented graphically for a tube with 500 scanning lines in the picture with a picture diagonal of 635 mm (25 in.), 90° deflection, a screen voltage of 25 kV, a cathode temperature of 1160°K (characteristic of an oxide cathode), and a cathode emission of 1 A/cm^2. Figure 4.3 shows the upper limit to the beam current established by thermal spread and final-lens aberration, Fig. 4.8 shows the limit determined by space charge in the drift region, while Figs. 4.4–4.7 show the diameter of the emitting region of the cathode, the beam diameter at the final lens, the voltage ratio of the final lens, and the magnification of the final lens, respectively. Throughout, the quantity in question is plotted as a function of the clear diameter of the final lens, with the focal length of the beam-forming region as parameter. The focal length of the beam-forming region is determined by the dimensioning and spacing of the electrodes in the beam-forming region and by their potentials.

Figures 4.3 and 4.8 indicate the advantage of a large clear diameter of the final lens for achieving high beam current. It must be remembered, however, that the analysis of this chapter is limited to the undeflected beam. With deflection (particularly, deflection through large angles), both the power needed for deflection and the aberrations contributed by the deflecting field increase with the beam diameter and, for three-beam tubes, also increase with the beam spacing. The beam spacing, in turn, is closely related to the maximum permissible clear diameter of the focusing elements of the individual beam. Thus, considerations of deflection power economy, the need for deflection aberration compensation, and the achievement of adequate brightness and color purity in the peripheral parts of the picture may dictate the choice of relatively small clear diameters in the final lens elements.

The mathematical derivation of the beam-current limitations just described is now given.

The screen brightness (luminance) B, which is the quantity of primary interest, is given by the expression

$$B = (T_g E_{conv})[T(I_s V_s / HW)][1 - (t_b/t)] \tag{4.1}$$

Here T_g is the optical transmission of the faceplate, E_{conv} is the conversion efficiency of the phosphor screen, T is the fraction of the beam electrons which strike the phosphor (equal to the mask transmission in a shadow-mask tube with full phosphor coverage of the screen), I_s is the beam current, V_s is the screen voltage, H and W are the height and width of the picture, and t_b/t is the fraction of the total time during which the beam is blanked or overscanned. The conversion efficiency is measured in light flux emitted in a forward direction per unit incident beam power and is assumed to take account both of electron energy absorption within the aluminum film (or any other film deposited on the phosphor to improve contrast) and of the luminance gain resulting from reflection at the aluminum film. In three-beam tubes, the screen brightness becomes, of course, the sum of three terms similar to that given by Eq. (4.1), each with its own values of I_s and E_{conv}.

In Eq. (4.1), applied to a tube of given design, all the factors with the exception of the beam current I_s can be regarded as either known or prescribed. Thus, to maximize the screen brightness, we must seek to maximize the beam current, subject to constraints established by the required resolution, the cathode temperature and emission capability, the anode voltage, the total deflection angle, the picture dimensions, and the neck diameter of the tube.

It appears reasonable to require that the vertical dimension of the spot be at most equal to the center-to-center separation of the scanning lines, h. With this choice, the scanning lines can be resolved and the spot should not deteriorate vertical resolution appreciably. More precisely, if the spot distribution is gaussian, we shall require

$$I(r) = \exp(-4r^2/h^2) \tag{4.2}$$

and correspondingly for the intensity distribution across a scanning line,

$$I(y) = \exp(-4y^2/h^2) \tag{4.3}$$

This means that for an isolated scanning line the intensity at the effective "edge" of the scanning line, $y = h/2$, has dropped to a fraction 0.368 of the intensity at the line center. The intensity in a uniform field will then appear modulated at the spatial frequency of the scanning lines with a modulation amplitude, $\exp(-\pi^2/4) = 0.085$. For a 525-line, 30-field/sec picture with a 4:3 aspect ratio, the spatial frequency f_s corresponding to a cutoff frequency $f_c = 4$ MHz is

$$f_s = \frac{f_c}{v} = \frac{f_c(1 - t_b/t)}{(525)^2 \cdot 30 \cdot \frac{4}{3}h} = \frac{0.29}{h} \qquad \text{for} \quad \frac{t_b}{t} = 0.2 \tag{4.4}$$

Here v is the scanning velocity. The sine-wave response (modulation transfer function) for a gaussian spot with the diameter h becomes, at this frequency,

$$\exp[(-\pi^2/4)f_s^2h^2] = 0.81 \tag{4.5}$$

Thus, a circularly symmetric spot of diameter (as defined above), equal to the center-to-center line separation h, should not materially reduce the horizontal resolution or sharpness of the reproduced picture.

A circularly symmetric spot is a reasonable choice for a cathode-ray tube with a uniform screen as well as for a shadow-mask tube with a dot screen. For tubes with vertical line screens and especially with beam-index tubes, a spot contracted in the horizontal direction (perpendicular to the phosphor lines) is desirable. Thus, in beam-index tubes, three adjoining phosphor lines constitute the basic period of the screen structure in the horizontal direction and should hence be of the same order as the center-to-center scanning line separation h; the achievement of saturated colors demands, on the other hand, that the horizontal width of the scanning spot should not exceed a single-color line, or be approximately $h/3$.

The spot size is one of the chief determining factors for an upper limit to the beam current I_s. Other determining factors for the upper limit to the beam current are the cathode emission density j_0, the final-anode voltage V (which, in some tubes, differs from the screen voltage V_s), the effective cathode temperature T, the electron-optical aberrations of the imaging systems which form the spot, and eventually, space-charge spreading of the beam. Variations in spot size and registration with deflection can impose further limits, which will not be considered here.

The basic electron-optical system for forming the spot is illustrated in Fig. 4.1. The electrons emitted by the cathode are accelerated and con-

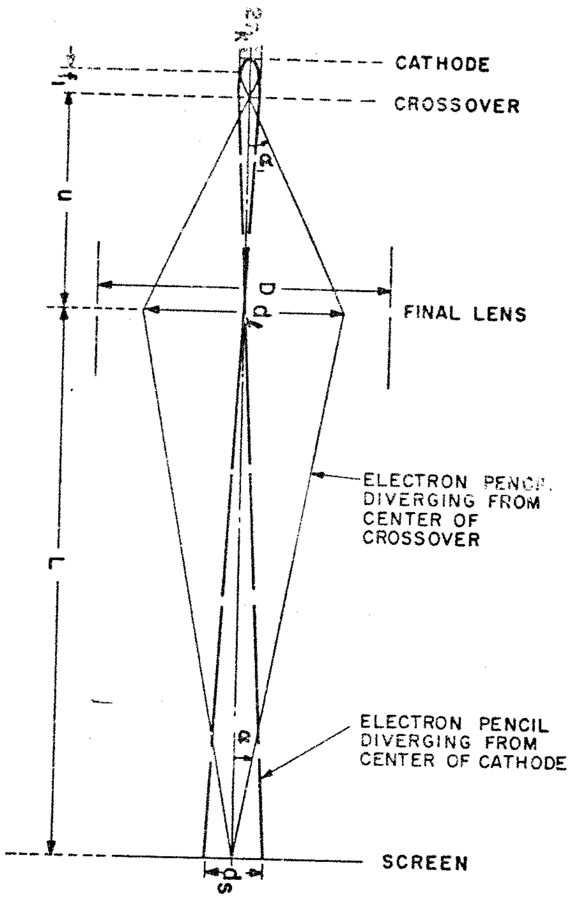

Fig. 4.1. Schematic diagram of electron-optical system forming the spot on the screen.

verge to (or appear to diverge from) a "crossover," which, in turn, is imaged electron optically on the screen to form the spot. The crossover is the point at which the "principal rays," or the electrons leaving points of the cathode with zero initial velocity, intersect the axis of the system. We can assign to the electric fields forming the crossover, or to the "first lens" of the electron gun, a focal length f_1, defined by

$$f_1 = r_k/\sin \alpha_1 \tag{4.6}$$

Here α_1 is the angle of inclination at the crossover of a principal ray leaving a cathode point a distance r_k from the axis. The crossover is at object distance u from the object-side principal plane of the "second lens" or "final lens" of the gun, and the screen lies at image distance L from the image-side principal plane of the same lens. If the potential at the crossover is V_1, the magnification of the spot with respect to the crossover is given by

$$M = L/u(V_1/V)^{1/2} = \sin \alpha_1/\sin \alpha (V_1/V)^{1/2} \tag{4.7}$$

Here α and V are the angle of convergence and potential, respectively, at the spot.

The preceding description is always valid, irrespective of the complexity of the imaging system of the gun, provided that axial symmetry is maintained. It is true that the parameters f_1, u, L, and M are constants only for small values of α and the corresponding quantities α_1 and r_k. For larger angles of inclination it is necessary to take account of the "spherical aberration" or aperture defect of the electron-optical system. Analytically it is then still convenient to treat the crossover as aberration-free and to include the aberration of the first lens in the aberration of the final lens system.

D. B. Langmuir (*33*) has shown that for an aberration-free system with a cathode with uniform emission current density j_0 on an emitting area of radius r_k, the current distribution at the crossover is given by

$$j_1(r_1) = j_0 \sin^2 \alpha_1 \frac{eV_1}{kT} \exp\left(-\frac{eV_1}{kT}\frac{r_1^2}{f_1^2}\right) \tag{4.8}$$

Here e is the magnitude of the electron charge.

Similarly, the current in the spot is given by

$$j(r) = j_0 \sin^2 \alpha \frac{eV}{kT} \exp\left(-\frac{eV}{kT}\frac{u^2}{L^2 f_1^2} r^2\right) \tag{4.9}$$

It will be noted that the emitting area of the cathode does not affect the current distribution in either case, but only the magnitude of the current, through the factors $\sin^2 \alpha_1$ and $\sin^2 \alpha$, respectively. Every point of the cathode contributes an identical gaussian distribution with a $(1/e)$ radius $f_1(kT/eV_1)^{1/2}$ or $f_1(L/u)(kT/eV)^{1/2}$, respectively.

The quantity $\sin \alpha$ is related to the emitting radius r_k of the cathode by Eqs. (4.6) and (4.7) only if there are no current-limiting apertures in the system. However, Eq. (4.9), with α signifying the actual convergence half-angle of the electron pencils at the screen, is still completely valid if the limiting aperture is in the plane of a cathode image. The coefficient $j_0 \sin^2 \alpha(eV/kT)$ represents the current density at the center of the spot formed by the aberration-free system even if the last-mentioned condition is not satisfied.

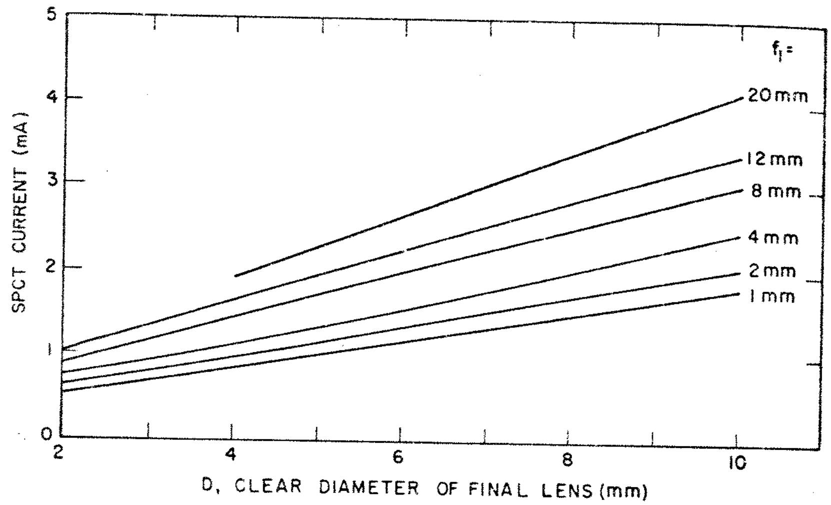

Fig. 4.8. Upper limit to spot current determined by space-charge spreading in drift region (for beam diameters at final lens given by Fig. 4.5).

charge imposes, for large values of the beam-forming system focal length f_1 and the clear diameter of the final lens D, a lower limit on the spot current than is imposed by the initial velocity distribution. This would no longer be true if the maximum emission density of the cathode were reduced to 0.5 A/cm^2. This would reduce the ordinates in Fig. 4.3 by a factor 2, but would have no effect on Fig. 4.8. The closeness of the space-charge limit to the thermal limit indicates, in any case, that proper consideration of space charge would lead to stronger final lenses at high currents and to upper limits to the spot current which are lower than those shown in Fig. 4.3.

4.2 Three-Beam Tubes

The screen-brightness considerations presented so far apply quite generally, and in particular also to black-and-white viewing tubes. In color viewing tubes employing three beams, there are two additional factors which reduce the attainable screen brightness. One of these, which will not be considered further at this point, is the replacement of the single white phosphor in the black-and-white viewing tubes by three red, green, and blue primary-color phosphors which do not possess as high an average conversion efficiency as the white phosphor. The other factor is the need to introduce some kind of mask to stop beam electrons which would otherwise be incident on the "wrong" phosphors. The screen brightness is then reduced by a factor equal to the electron transmission of the mask. As will be seen in Chapter 6, electric fields at the mask may be used to reduce the diameters of the electron pencils transmitted by the mask apertures, permitting greatly increased mask transmission. In the simplest "focus-mask" tube designs the requirement of lower anode voltage and unfavorable contrast properties offset the gain in mask transmission.

The usual shadow mask (Fig. 4.9) has a hexagonal array of round apertures of diameter B and center-to-center separation a. The three-component electron guns form a triangular "delta" arrangement, so that the beam-cross-section centers in the deflection plane have a common separation s from the tube axis. We shall call the beam diameters in the deflection plane d_l, since they differ but slightly from the beam diameters at the principal plane of the final lens. The beam cross sections in the deflection plane, located a distance p from the mask, project onto the screen, a distance q from the mask, a set of electron spots. This set, ideally, coincides with an array of round phosphor dots having diameters D_s and covers the screen surface.

The condition that the dots corresponding to a single aperture form a contiguous triad may be written

$$D_s/\sqrt{3} = s(q/p) \tag{4.28}$$

and the further condition that the dot triads fit together or "nest" so as to

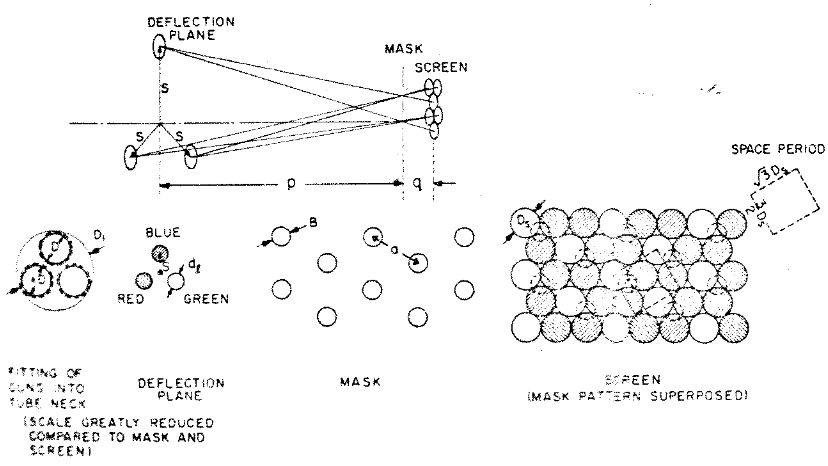

Fig. 4.9. Shadow-mask tube with hexagonal dot array; beam cross sections in deflection plane, mask pattern, and screen pattern.

uniformly cover the screen becomes

$$[(p + q)/p]a = \sqrt{3}D_s \tag{4.29}$$

Finally, the condition that the electron spots equal the phosphor dots in size determines the maximum value of the aperture diameter B consistent with the attainment of color purity,

$$[(p + q)/p]B + (q/p)\, d_l = D_s \tag{4.30}$$

The maximum value T_m of the mask transmission,

$$T = (\pi/2\sqrt{3})B^2/a^2 \tag{4.31}$$

is obtained by substituting the value of B from Eq. (4.30) in Eq. (4.31),

$$T_m = (\pi/6\sqrt{3})(1 - d_l/s\sqrt{3})^2 \tag{4.32}$$

Let the inner diameter of the gun neck be D_i, the external diameter of the individual component gun be kD $(k > 1)$, and the internal clear diameter of the final lens of the gun be D. Then, if the gun dimensions are made as large as possible (to minimize the spherical aberration), we have

$$D_i = kD(1 + 2/\sqrt{3}) = 2.154kD, \qquad kD = s\sqrt{3} \tag{4.33}$$

Hence, we can write in place of Eq. (4.32),

$$T_m = 0.302(1 - d_l/kD)^2 \tag{4.34}$$

Since the current reaching the screen* is proportional to d_l^2T, the optimum choice of d_l in the absence of spherical aberration would be $kD/2$, leading to a mask transmission T_m of only 0.075. In practice, as shown in Fig. 4.5, spherical aberration limits d_l to smaller values, ranging from $0.15D$ to $0.4D$. For example, for $d_l/kD = 0.3$, we would obtain $T_m = 0.15$. Shadow-mask tubes with screens fully covered by phosphor dots have generally employed masks with transmissions of this order. The requirement of providing printing tolerances and of accounting for imperfect registration between electron spots and dots tends to reduce the permissible mask transmission.

In the negative-tolerance matrix-screen shadow-mask tubes described in Chapter 5, Section 5.5, the phosphor dots are reduced in size and surrounded by a black matrix, diminishing the reflectivity of the viewing screen for ambient light. Since, here, the electron spots projected on the screen are larger than the phosphor dots, the dot size, along with the mask transmission, limits the efficiency with which the beam current delivered by any one gun is utilized for the excitation of the phosphor of the correspond-

* For uniform current density in the beam!

ing color. This efficiency can be made larger, however, than the mask transmission in shadow-mask tubes with nonmatrix screens. A typical value for the current utilization efficiency of negative-matrix tubes distributed in 1973 is 0.23. In such tubes, with a measured phosphor conversion efficiency of 40 lm/ W(when the three beam currents are adjusted to produce a black-and-white image) and a faceplate transmission of 0.85, the efficiency with which the beam power is utilized for useful light production becomes 8 lm/W.

By comparison, typical values of phosphor screen conversion efficiencies of black-and-white tubes of the same period are found to be 46 lm/W. With 78% faceplate transmission, the efficiency with which the beam power is utilized for the production of useful light becomes here 36 lm/W.

Compare now shadow-mask tubes and black-and-white tubes with identical cathode emission density, neck and bulb dimensions, and screen voltage. The maximum clear diameter of the final lens of the black-and-white tube gun can now be made 2.2 times as large as the clear diameter of one of the three-component, color-tube guns. Correspondingly, according to Eq. (4.24), the maximum current for the single gun of the black-and-white tube would be 2.5 times as large as that for one of the three-component guns of the color tube. If equal currents in the three-color-tube guns produce white, we would expect for the ratio of the maximum screen brightness of the color tube to that of the black-and-white tube,

$$\frac{\text{max. screen brightness for color tube}}{\text{max. screen brightness for black-white tube}} = \frac{8 \cdot 3}{36 \cdot 2.5} = 0.27$$

Since the beam power, and hence the screen brightness, is proportional to the square of the screen voltage, a negative-matrix shadow-mask tube operated at 27.5 kV should be capable of the same screen brightness as a black-and-white tube of similar dimensions operated at 14.3 kV.

With an in-line gun, slit mask and continuous-line phosphor screen (Fig. 4.10), we must write

$$D_s = s(q/p), \qquad [(p + q)/p]a = 3D_s \tag{4.35}$$

$$[(p + q)/p]B + (q/p)d_1 = D_s$$

Here D_s, s, a, B, and d_1 all denote dimensions in a direction transverse to the slits and phosphor lines.

With

$$T = (1 - c)B/a \tag{4.36}$$

we find

$$T = \tfrac{1}{3}(1 - c)(1 - d_1/s) \tag{4.37}$$

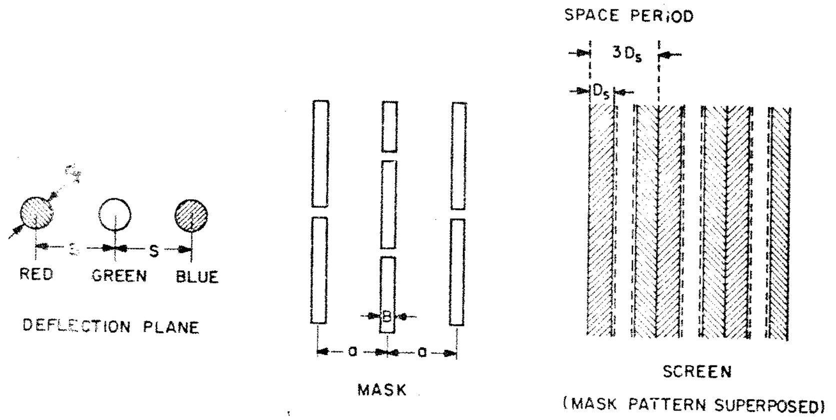

Fig. 4.10. Shadow-mask tube with slit mask and in-line gun: beam cross sections in deflection plane, mask pattern, and screen pattern.

Here c is the ratio of the bridge length to the slit length in the mask, which is a small quantity for masks having a bridge as shown in Fig. 4.10 and zero for masks with a continuous slit as in the Trinitron (*24, 39*). The aberrations of a large, single final lens, traversed at its center by the three beams, should permit a larger value of D, and hence of d_l, than in triple guns with separate focusing systems. The relation between D_i and s, corresponding to Eq. (4.33) for the delta system, is $D_i = 3s$.

4.3 Single-Gun Tubes

In all single-gun systems there is a loss in brightness resulting from the fact that the beam can be on any one of the three phosphors for at most one-third of the scanning time. Even a duty cycle of one-third will give high color purity only with the employment, for example, of rectangular waves to switch the beam from one phosphor to the next. The high subcarrier frequency makes the generation of such switching waves commonly impractical. Thus, either sinusoidal switching or switching waves formed by a few harmonics are employed. Color purity then demands a duty cycle substantially less than one-third. A fraction of this intensity loss (as compared with three-beam systems) may be made up by the use of guns with final lenses of larger clear diameter than is possible with three guns.

As compared with a black-and-white viewing tube of the same dimensions and anode voltage, the screen brightness (at equal screen voltage) is reduced by the indicated duty cycle, the ratio of the average phosphor