5.1 Electron Tube Fundamentals .................................. 353
Introduction • Cathodes and Thermionic Emission
 • Thermionic Diode • Grid Tubes • Vacuum Tube Specifications • Biasing Methods to Establish a DC Operating Point • Power Amplifier Classes • Power-Output Calculation

5.2 Power Grid Tubes ........................................... 367
Introduction • Vacuum Tube Design • High Frequency Operating Limits

5.3 Neutralization Techniques ................................. 387
Introduction • Circuit Analysis • Self-Neutralizing Frequency • Neutralization Adjustment

5.4 Amplifier Systems ........................................... 397
Introduction • Principles of RF Power Amplification • Cavity Amplifier Systems

5.5 Image Capture Devices .................................... 424
Introduction • The Plumbicon • Operating Characteristics

5.6 CRT Devices and Displays ............................... 430
Introduction • Basic Operating System • Color CRT Devices

5.7 Projection Systems ........................................... 451
Introduction • Projection System Fundamentals • Projection Display Systems • Operational Considerations

5.1 Electron Tube Fundamentals

Clifford D. Ferris
University of Wyoming

Jerry C. Whitaker
Editor-in-Chief

Steve Epstein
Video Systems Magazine

5.1.1 Introduction

Although widely used in many electronic applications until the 1960s, the traditional electron vacuum tube has now largely been replaced by semiconductor devices. Notable exceptions include cathode-ray tubes used in oscilloscopes and video displays, power grid tubes used in radio and television broadcast transmitters, and various classes of vacuum microwave devices. This chapter presents an introduction to the operation of electron tubes. Subsequent chapters are concerned with specific vacuum devices and their applications.

The elementary vacuum-tube structure is illustrated in Fig. 5.1. An evacuated volume is enclosed by an envelope, which may be glass, metal, ceramic, or a combination of these materials. Metallic electrodes
contained within the envelope form the working device, and external connections are made via metal wires that pass through vacuum seals in the envelope. The three basic internal structures are the cathode (K), the anode (A) (often called the plate), which are solid surfaces, and a single open wire-mesh structure that is the grid. Under typical operating conditions, the anode is held at a more positive potential than the cathode, and the grid is held at a voltage that is negative relative to the cathode. In this regard, operation is analogous to that of an n-channel junction field-effect transistor (JFET). The cathode, grid, and plate are analogous to the source, gate, and drain, respectively.

The function of the cathode is to provide a source of free electrons, and this is accomplished by heating the cathode to a high temperature, a process known as thermionic emission. Because the electrons are negatively charged and the cathode is lower in potential than the plate, the emitted electrons are accelerated toward the positively charged plate. The electron flow from the cathode to the plate can be controlled by adjusting the electric potential (grid-to-cathode voltage) applied between the grid and cathode. If the grid is made sufficiently negative relative to the cathode, the electron flow to the plate can be cut off completely. As the grid becomes less negative, electrons pass through the grid structure to the plate. The grid potential, however, must not increase to the point that it collects electrons in place of the plate. The grid is composed of fine wires that are not designed to support high current and is easily damaged by any significant electron collection.

5.1.2 Cathodes and Thermionic Emission

There are two fundamental cathode structures, directly heated (filament) and indirectly heated. The filament type of cathode is similar to the tungsten filament in an incandescent light bulb. To increase the efficiency of electron emission, thorium is added to the tungsten (1 or 2% by weight of thorium oxide), thus forming a thoriated tungsten filament (after an activation process that converts the thorium oxide to metallic thorium). The actual mechanical configuration or geometry of the filament varies widely among vacuum tubes depending on the output-power rating of the vacuum tube. Typically, the filament (or heater, as it is often called) is supported by the lead wires that pass through the envelope.

Figure 5.2 represents a vacuum diode that contains a directly heated cathode (filament). The filament supply voltage may be either AC or DC (DC is shown), and historically is called the A supply. Depending on the service for which a vacuum tube has been designed, filament (heater) supply voltages typically range from 2.5 to 12.6 V. Filament currents vary from as low as 150 mA to hundreds of amperes. A DC potential (the plate supply voltage) is required between the plate and cathode. Under many conditions, the cathode is at 0 V DC so that the plate-to-cathode voltage is the same as the plate voltage.
The thermionic-emission current $I_{th}$, in amperes, generated by a heated filament is described by the Richardson–Dushman equation

$$I_{th} = A D_0 T^2 \exp(-b_0 / T)$$

(5.1)

where

- $A$ = surface (emitting) area of the filament, m$^2$
- $D_0$ = material constant for a given filament type, A/m$^2$-K$^2$
- $T$ = absolute temperature, K
- $b_0 = 11,600 E_w$, K
- $E_w$ = the work function of the emitting surface, eV

Some typical values for thoriated tungsten are $D_0 = 3.0 \times 10^4$ A/m$^2$-K$^2$, $T = 1900$ K, $b_0 = 30,500$ K, and $E_w = 2.63$ eV. Cathode efficiency is defined in terms of the number of amperes (or milliamperes) produced per square meter of emitting surface per watt of applied filament power. For a thoriated tungsten filament with plate voltages in the range from 0.75 to 5 kV, emission efficiency ranges from 50 to 1000 A/m$^2$/W. Thoriated tungsten filaments are normally incorporated into power grid tubes that generate output power levels from a few hundred watts to thousands of kilowatts. These devices are used in high-power radio and television broadcasting service and in industrial applications such as induction heating.

Oxide-coated cathodes that are indirectly heated are more efficient ($E_w \approx 1$ eV) than thoriated tungsten filaments, but they cannot sustain the high-voltage high-current operating conditions associated with power grid tubes. When the maximum output power required is below a few hundred watts, they are the device of choice and find application in low-power high-frequency transmitters. They are also widely used by audiophiles who feel that vacuum-tube power amplifiers produce a quality of sound output that is superior to that produced by semiconductor devices.

A typical geometry for an oxide-coated cathode is a nickel or Konal-metal sleeve or cylinder that surrounds a heating filament (the heater). The heater wire is usually tungsten or a tungsten-molybdenum alloy that is coated with a ceramic insulating material (alundum, aluminum oxide or beryllium oxide) to isolate the heater electrically from the cathode sleeve. The outside coating of the cathode is applied as a mixture of carbonates that are subsequently converted to oxides during final processing of the electron tube. Barium and strontium oxides are widely used. Typical heater temperatures are on the order of 1000°C, which produce an orange to yellow-orange glow from the heater. Cathode temperatures are on the order of 850°C. Maximum heater-power requirements are on the order of 12 W for oxide-coated cathode devices, whereas thoriated tungsten filaments may require many kilowatts of heater input power.

Figure 5.3 illustrates two of many possible heater configurations. The twisted configuration (Fig. 5.3(a)) and the folded linear arrangement (Fig. 5.3(b)) are commonly used. The ceramic oxide coating...
coating that is applied to the wires prevents electrical short circuits between the wire segments. The heater is inserted into the cathode sleeve, and the ceramic coating applied to the heater surface insulates the heater electrically from the cathode sleeve. This ceramic layer cannot sustain a high potential difference between the heater and the cathode, and 200 V is a typical figure for the maximum heater-cathode voltage. For this reason, indirectly heated cathodes are not appropriate for power grid tube service.

Before we move on to specific vacuum-tube configurations, a final comment is required concerning electron emission from a cathode. To simplify mathematical derivations, it is often assumed that all of the electrons are emitted with zero initial velocity. If some of the electrons have initial velocities, a negative potential gradient is established at the cathode, which then produces a potential-energy barrier and corresponding space-charge region. Electrons emitted with very small initial velocities are then repelled toward the cathode. Equilibrium is established between emission of electrons and this space-charge region. In effect, the electrons that are accelerated toward the plate (the emission current) are drawn from the space-charge region.

5.1.3 Thermionic Diode

The fundamental thermionic diode has been illustrated in Fig. 5.2, and Fig. 5.4 illustrates the representation of a vacuum diode that has an oxide cathode. The plate current \( I_P \) vs plate-to-cathode voltage \( V_P \) description of diode operation is provided by the Langmuir–Child law that states that the plate current is proportional to the plate voltage raised to the three-halves power.

\[
I_P = k_0 V_P^{1.5}
\]

The proportionality constant \( k_0 \) depends on the physical dimensions of the plate and cathode, their respective geometries, and the ratio of electronic charge to mass. It is sometimes called the permeance. It must be noted that the plate current cannot exceed the thermionic emission current, which provides an upper limit to the value of \( I_P \) in Eq. (5.2). A typical characteristic curve is presented in Fig. 5.5, in which \( k_0 = 8.94 \times 10^{-6} \) (A/V^{1.5}).

Vacuum diodes are used as circuit elements in the same manner as modern semiconductor diodes. They may serve in rectifier circuits, and as voltage clippers, clamps, and DC restorers for high-voltage applications. From a circuit viewpoint, the only difference is the requirement of a heater voltage when a thermionic diode is used. The forward-voltage drop associated with vacuum diodes, however, is considerably larger than the fraction of a volt associated with a silicon device. It may range from a few volts to more than 100 V depending on the device, the plate current, and the plate-to-cathode voltage (plate voltage).

5.1.4 Grid Tubes

To achieve voltage and power amplification, or to produce electronic oscillation, at least one additional electrode must be added to the basic thermionic diode. This electrode is a control grid. As will be discussed subsequently, two additional grids may be added to produce specific operating characteristics of the grid tube. Figure 5.6 illustrates four device configurations. Depending on the maximum power dissipation of the device, either of the two cathode types shown may be used. For graphical convenience, the heater connections have been omitted in Fig. 5.6(a) to Fig. 5.6(d).
FIGURE 5.5  Vacuum diode characteristic for $k_p = 8.94 \times 10^{-6} \text{ (A/V}^{1.5})$.

FIGURE 5.6  Four standard grid tube representations: (a) triode, (b) tetrode, (c) pentode, (d) beam power, (e) cathode types.
Physical geometries for grid tubes vary widely depending on operating frequency and power dissipated. Many devices use a modified concentric cylinder configuration as indicated in Fig. 5.7. The electrode structures are supported by stiff metal posts anchored at the base (and sometimes the top) of the device. Ceramic or mica wafer spacers (not shown in the diagram) hold the electrode structure rigidly within the enclosing envelope. The plate is frequently equipped with heat radiating fins. The grid is often an open helix constructed from fine wire, although other structures resembling window screening are used in some applications.

Parallel planar geometries are used in some high-frequency triodes. A planar cathode is placed at the base of the tube, and a heavy planar plate is anchored at the top of the tube, with a flat-mesh control grid inserted between the cathode and plate. Close spacing between the plate and cathode is obtained in this manner so that electron-transit time from the cathode to the plate is minimized.

Nickel and various metals and alloys are used for the tube electrodes other than the heater. The plate is frequently blackened to improve heat radiation. In low-power grid tubes, the electrical connections are usually made through the base of the device. For low-power tubes there are various standard base-pin and corresponding socket configurations similar to those used for standard television picture tubes. In high-power and high-frequency tubes, the heater connections are usually made at the tube base, but the plate connection is at the top of the tube, and the grid connections are brought out as wires or rings along the side of the device. Power grid tubes are described in a separate chapter in this handbook.

By general convention, the cathode potential is assumed to be close to or at zero volts. Thus, references to grid volts, screen volts, and plate volts (rather than grid-to-cathode, screen-to-cathode, or plate-to-cathode) are standard in vacuum-tube literature. In low-power tubes, the control grid (G in Fig. 5.6) is operated at either zero volts or some negative voltage. The plate voltage is positive, the screen grid voltage (SG in Fig. 5.6) is positive but lower than the plate voltage, and the suppressor grid voltage (SUPG in Fig. 5.6) is usually at the same potential as the cathode. In many tubes, the suppressor grid is connected internally to the cathode.

The signal processing attributes of vacuum tubes are characterized by three primary parameters, amplification factor $\mu$ (mu), dynamic transconductance $g_m$, and dynamic plate resistance $r_p$, which are defined as follows:

$$\mu = \frac{\partial v_p}{\partial v_g}; \quad I_P \text{ and } V_{SG}, \text{ held constant as applicable}$$

$$g_m = \frac{\partial i_p}{\partial v_g}; \quad V_P \text{ and } V_{SG}, \text{ held constant as applicable}$$

$$r_p = \frac{\partial v_p}{\partial i_p}; \quad V_G \text{ and } V_{SG}, \text{ held constant as applicable}$$

$$\mu = g_m r_p$$

Note that lowercase letters have been used to define signal parameters, and uppercase letters have been used to define DC values. Thus, $\mu$, $g_m$, and $r_p$ are defined for a fixed DC operating point (Q point) of the corresponding device. Additional signal parameters include the direct interelectrode capacitances, which are a function of the geometry of a specific tube type. These capacitances are generally a few picofarads at most in low-power grid tubes and include, as applicable: control grid to plate; control grid to heater, cathode, screen grid, suppressor grid.

The input resistance presented between the grid and cathode of a vacuum tube is extremely high and is generally assumed to be infinite at low frequencies. The input impedance, however, is not infinite because of the influence of the interelectrode capacitances. Input impedance is thus a function of frequency, and it depends on the tube parameters $\mu$, $g_m$, and $r_p$ as well. Reactive elements associated with the load also influence input impedance. If $C_{gk}$ is the interelectrode capacitance between the grid and cathode, $C_{pg}$ is the capacitance between the plate and grid, $R_l$ is the effective AC load (no reactive/susceptive components), and
\( \omega \) is the radian frequency of operation for an amplifier tube, then input admittance \( (Y_i) \) is, approximately, for triode

\[
Y_i \approx j \omega \{ C_{gk} + \left[ \frac{\mu R_l}{(r_p + R_l)C_{pg}} \right] \}
\]  

(5.3a)

and tetrode and pentode

\[
Y_i \approx j \omega C_{gk}
\]  

(5.3b)

At very high frequencies, lead inductances and stray capacitances associated with the tube itself and its external connections modify the expressions shown.

With the exception of cathode emissivity, the operating parameters of vacuum tubes depend solely on geometry and the placement of the electrodes within an evacuated envelope. If mechanical tolerances are rigidly held during manufacture, then there is very little variation in operating parameters across vacuum tubes of the same type number. Cathode manufacture can also be controlled to high tolerances. For this reason, when an operating point is specified, single numbers define \( \mu, g_m, \) and \( r_p \) in vacuum tubes. The designer does not have to contend with the equivalent of beta spread in bipolar-junction transistors or transconductance spread in field-effect devices. To ensure reliable operation over the life of the tube, a barium getter is installed before the tube structure is sealed into an envelope and evacuated. After the tube has been evacuated and the final seal has been made, the getter is flashed (usually using a radio frequency induction-heating coil) and elemental barium is deposited on a portion of the inside of the envelope. The function of the getter is to trap gas molecules that are generated from the tube components with time. In this manner the integrity of the vacuum is preserved. Heater burnout or oxide-layer deterioration (in indirectly heated cathode tubes) are the usual causes of failure in vacuum tubes.

**Triode**

Historically, the triode shown in Fig. 5.6(a) was the first vacuum tube used for signal amplification. As noted in the introductory section (Fig. 5.1), the voltage applied between the grid (G) and the cathode (K) can be used to control the electron flow and, hence, the plate current. The dynamic operation of a triode is defined approximately by a modified form of the Langmuir–Child law as follows

\[
i_p = k_0 (v_g + v_p/\mu)^{1.5}
\]

(5.4)

The quantity \( k_0 \) is the permeance (as defined in the diode discussion) of the triode, and its value is specific to a particular triode type. The plate characteristic curves generated by Eq. (5.4), for a triode in which \( k_0 = 0.0012 \), are illustrated in Fig. 5.8. As the plate characteristic curves indicate, the triode is approximately a voltage-controlled voltage source.

![Triode plate characteristics.](image-url)
The value of $\mu$ in triodes typically ranges from about 4 (low-mu triode) to 120 (high-mu triode). Typical ranges for $r_p$ and $g_m$ are, respectively, from 800 $\Omega$ to 100 k$\Omega$, and from 1 to 5 mS.

**Tetrode**

In general, the triode exhibits a high direct capacitance (relatively speaking) between the control grid and the plate. To provide electrostatic shielding, a second grid (the screen grid) is added between the control grid and the plate. The screen grid must be configured geometrically to provide the amount of shielding desired, and arranged so that it does not appreciably affect electron flow. Insertion of the screen grid, however, does significantly change the plate characteristic curves, as is shown in Fig. 5.9. The negative-resistance region below 100 V may lead to unstable operation and care must be taken to assure that the plate voltage remains above this region. In the early days of radio, the negative-resistance of the tetrode was used in conjunction with tuned circuits to produce sinusoidal oscillators.

**Pentode**

To suppress the negative-resistance characteristic generated by the tetrode, a third coarse grid structure (the suppressor grid) is inserted between the screen grid and the plate. In this manner, the plate characteristic curves are smoothed and take the form shown in Fig. 5.10.

The dynamic operation of a pentode is described approximately by the relation

$$i_p = k_0(v_g + v_p/\mu_1 + v_g/\mu_2)^{1.5}$$

where $\mu_1$ is the amplification factor for the plate and control grid and $\mu_2$ is the amplification factor for the plate and screen grid. As the plate characteristic curves indicate, the pentode is approximately a voltage-controlled current source, and the curves are similar to those generated by an $n$-channel depletion-mode metal oxide semiconductor field effect transistor (MOSFET). The screen grid draws an appreciable current that may amount to as much as 20–30% of the plate current.
The value of $\mu$ is not normally provided on the specifications sheets for pentodes, and only the values for $g_m$ and $r_p$ are listed. Transconductance values are typically on the order of 5–6 mS, and plate resistance values vary from 300 k$\Omega$ to over 1 M$\Omega$.

Pentodes are sometimes operated in a triode connection, in which the screen grid is connected to the plate. In the 1950s, this mode of operation was popular in high-fidelity vacuum-tube audio power amplifiers.

**Beam Power Tube**

The beam power tube is a modified form of the tetrode or pentode that utilizes directed electron beams to produce a substantial increase in the output power. Beam-forming electrodes are used to shape the electron beam into the form of a fan (as indicated by the dotted lines in Fig. 5.11). An optional suppressor grid may be included. The control and screen grids are exactly aligned such that the electron beam is broken into dense fan-shaped layers or sheets. Beam power tubes are characterized by high-power output, high-power sensitivity, and high efficiency. The operating screen current ranges from approximately 5–10% of the plate current, as opposed to pentodes in which the screen current is on the order of 20–30% of the plate current.

![FIGURE 5.11 Cross-section of a beam power tetrode as viewed from the top.](image)

### 5.1.5 Vacuum Tube Specifications

Data sheets for vacuum tubes typically provide the information shown in Table 5.1 for a hypothetical low-power beam power tube. Some of the parameters and operating conditions specified in the table have yet to be addressed, and these matters are discussed subsequently.

### 5.1.6 Biasing Methods to Establish a DC Operating Point

As is the case with any electronic active device that is used in amplifier service, vacuum tubes must be biased for a DC operating point (Q point). Figure 5.12 illustrates this process graphically for a triode when the plate supply voltage is 400 V and the grid potential is $-8$ V. The DC load line is $R_L = (400 \text{ V}) / (80 \text{ mA}) = 5 \text{ k}\Omega$.

In fixed-bias operation, the control grid is supplied from a negative fixed-voltage bias supply (historically called the C supply). The plate is operated from a high-voltage positive supply. The circuit diagram for a resistor–capacitor–(RC–) coupled triode audio amplifier operating under fixed-bias conditions is shown in Fig. 5.13(a). The DC load line is achieved by using a 5-k$\Omega$ resistor in the plate circuit. The cathode is connected directly to ground, and the control grid is connected through a series resistor to the negative 8-V bias supply. The value of the grid resistor is generally not critical. It is inserted simply to isolate the applied input signal from the bias supply. Based on the Q-point location in Fig. 5.12, the plate current is approximately 28 mA and the voltage at the plate is approximately 260 V. Thus the Q-point parameters are: $V_C = -8 \text{ V}$, $V_P = 260 \text{ V}$, $I_P = 28 \text{ mA}$. Based on the data (not included here) used to generate the characteristic curves, $\mu = 16$, $g_m = 4.5 \text{ mS}$, and $r_p = 3550 \text{ }\Omega$.

Figure 5.13(b) illustrates the same audio amplifier configured for self-bias operation. The IR drop across a cathode resistor is used to place the cathode at the positive value of the grid potential above ground, whereas the grid is returned to ground through a resistor. Traditionally the grid resistor is called the grid-leak resistor, because its function is to bleed away from the grid any static charge that might build up.
TABLE 5.1 Operating Specifications for Hypothetical Beam Power Pentode Single-Tube Operation

<table>
<thead>
<tr>
<th>General:</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Heater voltage (AC/DC), V</td>
<td>6.3</td>
</tr>
<tr>
<td>Heater current, A</td>
<td>0.9</td>
</tr>
<tr>
<td>Direct interelectrode capacitances</td>
<td></td>
</tr>
<tr>
<td>Control grid to plate, pF</td>
<td>0.4</td>
</tr>
<tr>
<td>Control grid to cathode, pF</td>
<td>10</td>
</tr>
<tr>
<td>Plate to cathode, heater, screen, suppressor, pF</td>
<td>12</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Maximum Ratings:</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Plate voltage, V</td>
<td>400</td>
</tr>
<tr>
<td>Screen voltage, V</td>
<td>300</td>
</tr>
<tr>
<td>Plate dissipation, W</td>
<td>20</td>
</tr>
<tr>
<td>Screen-grid input, W</td>
<td>2.5</td>
</tr>
<tr>
<td>Peak heater-to-cathode voltage (positive or negative), V</td>
<td>180</td>
</tr>
</tbody>
</table>

**Typical Operating Conditions** (class A amplifier):

<table>
<thead>
<tr>
<th>Fixed bias</th>
<th>Cathode bias</th>
</tr>
</thead>
<tbody>
<tr>
<td>Plate supply voltage, V</td>
<td>350</td>
</tr>
<tr>
<td>Screen supply voltage, V</td>
<td>250</td>
</tr>
<tr>
<td>Control-grid voltage, V</td>
<td>−18</td>
</tr>
<tr>
<td>Cathode-bias resistor, Ω</td>
<td>—</td>
</tr>
<tr>
<td>Peak audio-frequency control-grid voltage, V</td>
<td>18</td>
</tr>
<tr>
<td>Zero-signal plate current, mA</td>
<td>54</td>
</tr>
<tr>
<td>Maximum-signal plate current, mA</td>
<td>66</td>
</tr>
<tr>
<td>Zero-signal screen-grid current, mA</td>
<td>2.5</td>
</tr>
<tr>
<td>Maximum-signal screen-grid current, mA</td>
<td>7</td>
</tr>
<tr>
<td>Plate resistance, Ω</td>
<td>33000</td>
</tr>
<tr>
<td>Transconductance, mS</td>
<td>5.2</td>
</tr>
<tr>
<td>Load resistance, Ω</td>
<td>4200</td>
</tr>
<tr>
<td>Total harmonic distortion, %</td>
<td>15</td>
</tr>
<tr>
<td>Maximum-signal power output, W</td>
<td>11</td>
</tr>
</tbody>
</table>

**Maximum Ratings** (Triode Connection: Screen Grid Connected to Plate):

| Plate voltage, V | 300 |
| Plate dissipation, W | 20 |
| Peak heater-to-cathode voltage (positive or negative), V | 180 |

**Typical Operating Conditions** (Class A Amplifier):

<table>
<thead>
<tr>
<th>Fixed bias</th>
<th>Cathode bias</th>
</tr>
</thead>
<tbody>
<tr>
<td>Plate supply voltage, V</td>
<td>250</td>
</tr>
<tr>
<td>Control-grid voltage, V</td>
<td>−20</td>
</tr>
<tr>
<td>Cathode-bias resistor, Ω</td>
<td>—</td>
</tr>
<tr>
<td>Peak audio-frequency control-grid voltage, V</td>
<td>20</td>
</tr>
<tr>
<td>Zero-signal plate current, mA</td>
<td>40</td>
</tr>
<tr>
<td>Maximum-signal plate current, mA</td>
<td>45</td>
</tr>
<tr>
<td>Plate resistance, Ω</td>
<td>1700</td>
</tr>
<tr>
<td>Amplification factor</td>
<td>8</td>
</tr>
<tr>
<td>Transconductance, mS</td>
<td>4.7</td>
</tr>
<tr>
<td>Load resistance, Ω</td>
<td>5000</td>
</tr>
<tr>
<td>Total harmonic distortion, %</td>
<td>5</td>
</tr>
<tr>
<td>Maximum-signal power output, W</td>
<td>1.4</td>
</tr>
</tbody>
</table>

on the grid structure. Since in normal operation the grid draws no current, the grid voltage rests at ground potential or zero volts. The cathode is biased above ground to provide the required value of grid-to-cathode potential. In the example case, if the cathode is at +8 V and the grid at 0 V, the effective grid-to-cathode voltage is −8 V. The value of the cathode resistor is calculated as \((8 \text{ V})/(28 \text{ mA}) = 286 \, \Omega\). The closest standard value is 270 Ω, as is shown in Fig. 5.13(b). To achieve maximum signal gain from the amplifier, the cathode resistor is bypassed by an appropriate value of capacitance.

A graphical approach may also be applied to evaluate the amplification properties of the circuits shown in Fig. 5.13. The AC load line \((R_L)\) for the circuits shown is the parallel combination of the 5-kΩ plate
resistor and the external signal load. Let us assume that \( R_l = 3.3 \, \text{k}\Omega \) and we now add it to Fig. 5.12 with the new plot shown in Fig. 5.14. The AC and DC load line intersect at the Q point. The amplifier gain is evaluated from the plot as follows. Apply to the grid a sinusoidal signal (\( \nu_i \)) with a peak value of 8 V. Thus, the grid voltage will swing from the Q-point value of \(-8\) V to \(0\) V and \(-16\) V, as indicated by the double-headed arrow. The corresponding plate voltages are indicated by the vertical arrows. When \( \nu_i = +8\) V, \( \nu_g = 0\) V and the plate voltage is \(\approx 190\) V; when \( \nu_i = -8\) V, \( \nu_g = -16\) V and the plate voltage

![Figure 5.12](image)

**FIGURE 5.12** Triode plate characteristics with superimposed DC load line for Q-point calculations.

![Figure 5.13](image)

**FIGURE 5.13** Triode audio amplifier circuits: (a) audio amplifier with fixed bias, (b) audio amplifier with cathode bias (self-bias).
is ≈320 V. Note that plate voltage decreases when grid voltage increases, thus the output signal is 180° out of phase with the input signal. The voltage gain is, thus, \( A_v = -(320 - 190)/16 = -8.1 \).

The output current has the same phase as the grid voltage, and the peak values are indicated by the horizontal arrows. When the grid voltage is 0 V, the plate current is \( \approx 49 \) mA, and when the grid voltage is \(-16\) V, the plate current is \( \approx 9 \) mA. Note that the output-voltage swing about the Q point is not quite symmetric. The Q-point value of the plate voltage is \( \approx 260 \) V. The positive swing in plate voltage is \( 320 - 260 = 60 \) V, and the negative swing is \( 260 - 190 = 70 \) V. This asymmetry is caused by the fact that the AC load line intersects the \(-16\)-V grid line at the onset of its curvature.

Based on a small-signal model of the triode, it can be shown analytically that the voltage gain of the amplifier configurations shown in Fig. 5.13 is

\[
A_v = -\mu / (1 + r_p / R_l) = -16 / (1 + 3550/3300) = -7.7
\]  

Note that there is a slight discrepancy between the voltage gain of \(-8.1\) found graphically and the calculated value of \(-7.7\). Some estimation errors occur when graphical analysis is used, and the solution is to use expanded graphical curves. The main source of error, however, is the use of a small-signal analysis result as predicted by Eq. (5.6) when the actual operation is a large-signal swing along the AC load line. Because vacuum tubes of the same type number are nearly identical and the plate characteristic curves are meaningful, in many cases more accurate results are obtained by graphical analysis than by small-signal calculations.

As is illustrated in Fig. 5.15, a DC load line is applied to the pentode plate characteristics in the same manner as for the triode. In the example shown, \( R_L = 1 \) kΩ and the Q-point values are \( V_p = 300 \) V, \( I_p = 200 \) mA, \( V_G = -9 \) V. Note for this device that positive-grid operation is possible and that nonlinearities occur for large negative grid voltage because the curves are not evenly spaced. An AC load is added to the pentode plate characteristics in the same manner as for the triode case, and the graphical analysis proceeds in the same manner as for the triode.

Two audio amplifier circuits (fixed bias and self-bias) are illustrated in Fig. 5.16. In the self-bias case, Fig. 5.16(b), the Q-point conditions have been assumed: \( I_p = 180 \) mA, \( V_G = -10 \) V, \( I_{SG} = 5 \) mA, \( V_{SG} = 350 \) V. The cathode resistor is, thus, \((10 \text{ V})/(180 \text{ mA}) = 55.6 \Omega\), and the closest standard value is 56 \( \Omega \). The screen resistor is \((500 \text{ V} - 350 \text{ V})/(5 \text{ mA}) = 30 \) kΩ, which is a standard value.

At low frequencies, the voltage gain of a pentode amplifier, such as the circuits shown in Fig. 5.16 is given by

\[
A_v = -g_{m}R_l
\]  

Copyright 2005 by Taylor & Francis Group
5.1.7 Power Amplifier Classes

Power amplifiers may be divided into various classes, designated by letters of the alphabet, according to how long the power grid tube conducts during one complete cycle of a sinusoidal input signal. Classes A, B, and C apply to analog operation, whereas additional designations (D–S) apply to pulsed operation. Pulsed-mode classes are addressed in other chapters.

In class A operation (sometimes designated as A1), the tube is biased as in Fig. 5.12 and Fig. 5.15, and conduction occurs for a full cycle. Because the tube is always on and drawing quiescent power \( P_Q = V_{QP} I_{QP} \), operating efficiency is low. In class B operation, the tube is biased such that it conducts for exactly one-half cycle. In class C operation, the tube is biased such that it conducts for less than one-half cycle but more than one-quarter of a cycle. Class C is the most efficient mode of these classes, and it is used in radio-telephony power amplifiers in which the associated high-Q circuits restore the missing portions of each cycle.

Class B operation is adopted for push-pull amplifier design in which the two tubes conduct during alternate half-cycles, thus reproducing the complete input signal. Harmonic distortion is introduced if...
either of the two devices fails to conduct for a full half-cycle. To avoid this eventuality, push–pull power amplifiers are frequently operated in a class AB mode in which both tubes conduct for slightly more than one-half cycle. Various designations, such as AB1 and AB2, have traditionally been applied to indicate the conduction time per cycle.

For single-tube operation, the maximum efficiency figures are 25% for class A, 50% for class B, and on the order of 70% for class C. In push-pull amplifier configurations, the maximum theoretical efficiency obtainable with transformer-coupled class A operation is 50%, and 78.5% in class B transformer-coupled or complementary-symmetry designs.

5.1.8 Power-Output Calculation

To complete this section, a sample power-output calculation is provided for the class A amplifier operation described by Fig. 5.14. The peak-to-peak swings of the output voltage and current along the AC load line are, respectively, $\Delta v_p = (320 - 190) = 130$ V, $\Delta i_p = (49 - 9) = 40$ mA. The output-signal power ($P_0$) is, thus

$$P_0 = \Delta v_p \Delta i_p / 8 = 0.125 \times 130 \times 0.04 \times 1000 = 650 \text{ mW}$$

In power triodes, output distortion is primarily generated by a second harmonic component, and the percentage distortion may be calculated from the empirical relation

$$\% \text{ distortion} = 100 \times [0.5(i_{p \text{ min}} + i_{p \text{ max}}) - I_{PQ}] / \Delta i_p \quad (5.8)$$

In our example, $I_{PQ} = 28$ mA = 0.028 A, $i_{p \text{ min}} = 9$ mA = 0.009 A, $i_{p \text{ max}} = 49$ mA = 0.049 A, $\Delta i_p = 40$ mA = 0.04 A, therefore

$$\% \text{ distortion} = 100 \times (0.5(0.058) - 0.028)/0.04 = 2.5\%$$

The quiescent (zero-signal) power ($P_Q$) is $260 \text{ V} \times 28 \text{ mA} = 7.28 \text{ W}$. Thus the percent efficiency is

$$\% \text{ efficiency} = 100 \times P_0 / P_Q = 100 \times 0.65/7.28 = 8.9\%$$

Defining Terms

**Cathode:** The heated electrode in a vacuum tube that emits electrons.

**Emissivity:** A measure of the degree of emission of thermionic electrons from a hot cathode or filament.

**Mu:** Dimensionless voltage amplification factor of a vacuum tube.

**Pentode:** Vacuum tube with five active electrodes: cathode, control grid, screen grid, suppressor grid, plate.

**Plate:** The positive electrode or anode in a vacuum tube.

**Tetrode:** Vacuum tube with four active electrodes: cathode, control grid, screen grid, plate.

**Thermionic:** Pertaining to thermal generation of electrons.

**Thoriated:** Pertains to a metal to which the element thorium has been added.

**Triode:** Vacuum tube with three active electrodes: cathode, control grid, plate.

References


Further Information

For additional information on the design, construction, and application of vacuum tubes to electronic circuits, the following books are recommended.


5.2 Power Grid Tubes

Jerry C. Whitaker

5.2.1 Introduction

A power-grid tube is a device using the flow of free electrons in a vacuum to produce useful work.\(^1\) It has an emitting surface (the cathode), one or more grids that control the flow of electrons, and an element that collects the electrons (the anode). As discussed in the previous chapter, power tubes can be separated into groups according to the number of electrodes (grids) they contain. The physical shape and location of the grids relative to the plate and cathode are the main factors that determine the amplification factor \( \mu \) and other parameters of the device. The physical size and types of material used to construct the individual elements determines the power capability of the tube. A wide variety of tube designs are available to commercial and industrial users. By far the most common are triodes and tetrodes.

5.2.2 Vacuum Tube Design

Any particular power vacuum tube may be designed to meet a number of operating parameters, the most important of which are usually high operating efficiency and high-gain/bandwidth properties. Above all, the tube must be reliable and provide long operating life. The design of a new power tube is a lengthy process that involves computer-aided calculations and modeling. The design engineers must examine a laundry list of items, including

- **Cooling**: how the tube will dissipate heat generated during normal operation. A high-performance tube is of little value if it will not provide long life in typical applications. Design questions include whether the tube will be air cooled or water cooled, the number of fins the device will have, and the thickness and spacing of the fins.
- **Electro-optics**: how the internal elements line up to achieve the desired performance. A careful analysis must be made of what happens to the electrons in their paths from the cathode to the anode, including the expected power gain of the tube.
- **Operational parameters**: what the typical interelectrode capacitances will be, and the manufacturing tolerances that can be expected. This analysis includes: spacing variations between elements within the tube, the types of materials used in construction, the long-term stability of the internal elements, and the effects of thermal cycling.

\(^1\) Portions of this chapter were adapted from Varian. 1982. *Care and Feeding of Power Grid Tubes*. Varian/Eimac, San Carlos, CA. Used with permission.
Device Cooling

The first factor that separates tube types is the method of cooling used: air, water, or vapor. Air-cooled tubes are common at power levels below 50 kW. A water-cooling system, although more complicated, is more effective than air cooling—by a factor of 5–10 or more—in transferring heat from the device. Air cooling at the 100-kW level is not normally used because it is virtually impossible to physically move sufficient air through the device (if the tube is to be of reasonable size) to keep the anode sufficiently cool. Vapor cooling provides an even more efficient method of cooling a power amplifier (PA) tube than water cooling. Naturally, the complexity of the external blowers, fans, ducts, plumbing, heat exchangers, and other hardware must be taken into consideration in the selection of a cooling method. Figure 5.17 shows how the choice of cooling method is related to anode dissipation.

Air Cooling

A typical air-cooling system for a transmitter is shown in Fig. 5.18. Cooling system performance for an air-cooled device is not necessarily related to airflow volume. The cooling capability of air is a function of its mass, not its volume. An appropriate airflow rate within the equipment is established by the manufacturer, resulting in a given resistance to air movement.

The altitude of operation is also a consideration in cooling system design. As altitude increases, the density (and cooling capability) of air decreases. To maintain the same cooling effectiveness, increased airflow must be provided.
**Water Cooling**

Water cooling is usually preferred over air cooling for power outputs above about 50 kW. Multiple grooves on the outside of the anode, in conjunction with a cylindrical jacket, force the cooling water to flow over the surface of the anode, as illustrated in Fig. 5.19.

Because the water is in contact with the outer surface of the anode, a high degree of purity must be maintained. A resistivity of 1 mΩ/cm (at 25°C) is typically specified by tube manufacturers. Circulating water can remove about 1 kW/cm² of effective internal anode area. In practice, the temperature of water leaving the tube must be limited to 70°C to prevent the possibility of spot boiling.

After leaving the anode, the heated water is passed through a heat exchanger where it is cooled to 30–40°C before being pumped back to the tube.

**Vapor-Phase Cooling**

Vapor cooling allows the permissible output temperature of the water to rise to the boiling point, giving higher cooling efficiency compared with water cooling. The benefits of vapor-phase cooling are the result of the physics of boiling water. Increasing the temperature of one gram of water from 40 to 70°C requires 30 cal of energy. However, transforming one gram of water at 100°C into steam vapor requires 540 cal. Thus, a vapor-phase cooling system permits essentially the same cooling capacity as water cooling, but with greatly reduced water flow. Viewed from another perspective, for the same water flow, the dissipation of the tube may be increased significantly (all other considerations being the same).

A typical vapor-phase cooling system is shown in Fig. 5.20. A tube incorporating a specially designed anode is immersed in a boiler filled with distilled water. When power is applied to the tube, anode dissipation heats the water to the boiling point, converting the water to steam vapor. The vapor passes to a condenser, where it gives up its energy and reverts to a liquid state. The condensate is then returned to the boiler, completing the cycle. Electric valves and interlocks are included in the system to provide for operating safety and maintenance. A vapor-phase cooling system for a transmitter with multiple PA tubes is shown in Fig. 5.21.

**Special Applications**

Power devices used for research applications must be designed for transient overloading, requiring special considerations with regard to cooling. Oil, heat pipes, refrigerants (such as Freon), and, where high-voltage holdoff is a problem, gases (such as sulfahexafluoride) are sometimes used to cool the anode of a power tube.

**Cathode Assembly**

The ultimate performance of any vacuum tube is determined by the accuracy of design and construction of the internal elements. The requirements for a successful tube include the ability to operate at high temperatures and withstand physical shock. Each element is critical to achieving this objective.

The cathode used in a power tube obtains the energy required for electron emission from heat. The cathode may be directly heated (filament type) or indirectly heated. The three types of emitting surfaces most commonly used are

- Thoriated tungsten
- Alkaline-earth oxides
- Tungsten barium aluminate-impregnated emitters
The thoriated tungsten and tungsten-impregnated cathodes are preferred in power tube applications because they are more tolerant to ion bombardment. The characteristics of the three emitting surfaces are summarized in Table 5.2.

A variety of materials may be used as a source of electrons in a vacuum tube. Certain combinations of materials are preferred, however, for reasons of performance and economics.

**Oxide Cathode**

The conventional production-type oxide cathode consists of a coating of barium and strontium oxides on a base metal such as nickel. Nickel alloys, in general, are stronger, tougher, and harder than most nonferrous alloys and many steels. The most important property of nickel alloys is their ability to retain strength and toughness at elevated temperatures. The oxide layer is formed by first coating a nickel structure (a can or disc) with a mixture of barium and strontium carbonates, suspended in a binder material. The mixture is approximately 60% barium carbonate and 40% strontium carbonate.

During vacuum processing of the tube, these elements are baked at high temperatures. As the binder is burned away, the carbonates are subsequently reduced to oxides. The cathode is then activated and will emit electrons.

An oxide cathode operates CW at 700–820 °C and is capable of an average emission density of 100–500 mA/cm². High-emission current capability is one of the main advantages of the oxide cathode. Other advantages include high peak emission for short pulses and low operating temperature. As shown in Table 5.2, peak emission of up to 20 A/cm² is possible from an oxide cathode. A typical device is shown in Fig. 5.22.

Although oxide-coated emitters provide more peak emission per watt of heating power than any other type, they are not without their drawbacks. Oxide emitters are more easily damaged or poisoned than other emitters, and also deteriorate more rapidly when subjected to bombardment of high-energy particles.

The oxide cathode material will evaporate during the life of the tube, causing free barium to migrate to other areas within the device. This evaporation can be minimized in the design stage by means of a high-efficiency cathode that runs as cool as possible but still is not emission limited at the desired heater
FIGURE 5.21 Vapour-phase cooling system for a 4-tube transmitter using a common waer supply.
TABLE 5.2  Characteristics of Common Thermonic Emitters

<table>
<thead>
<tr>
<th>Emitter</th>
<th>Heating Method(^a)</th>
<th>Operating Temperature, °C</th>
<th>Average Emission Density, A/cm²</th>
<th>Peak Emission Density, A/cm²</th>
</tr>
</thead>
<tbody>
<tr>
<td>Oxide</td>
<td>Direct and indirect</td>
<td>700–820</td>
<td>0.10–0.50</td>
<td>0.10–20</td>
</tr>
<tr>
<td>Thoriated tungsten</td>
<td>Direct</td>
<td>1600–1800</td>
<td>0.04–0.43</td>
<td>0.04–10</td>
</tr>
<tr>
<td>Impregnated tungsten</td>
<td>Direct and indirect</td>
<td>900–1175</td>
<td>0.5–8.0</td>
<td>0.5–12</td>
</tr>
</tbody>
</table>

\(^a\)Directly heated refers to a filament type cathode. (Source: Adapted from Fink, D. and Christiansen, D., eds. 1989. *Electronics Engineers’ Handbook.* McGraw-Hill, New York.)

Voltage. In the field, the heater voltage must not exceed the design value. An oxide cathode that is over-heated produces little, if any, additional emission. The life of the tube under such operation, however, is shortened significantly.

**Thoriated Tungsten Cathode**

The thoriated tungsten filament is another form of atomic-film emitter commonly used in power grid tubes. Tungsten is stronger than any other common metal at temperatures over 3500°F. The melting point of tungsten is 6170°F, higher than that of any other metal. The electrical conductivity of tungsten is approximately \(\frac{1}{3}\) that of copper, but much better than the conductivity of nickel, platinum, or iron-base alloys. The resistivity of tungsten in wire form is exploited in filament applications. The thoriated-tungsten filament (or cathode) is created in a high-temperature gaseous atmosphere to produce a layer of ditungsten carbide on the surface of the cathode element(s). Thorium is added to tungsten in the process of making tungsten wire. The thorium concentration is typically about 1.5%, in the form of thoria. By proper processing during vacuum pumping of the tube envelope, the metallic thorium is brought to the surface of the filament wire. The result is an increase in emission of approximately 1000 times over a conventional cathode.

At a typical operating temperature of 1600–1800°C, a thoriated tungsten filament will produce an average emission of 40–430 mA/cm². Peak current ranges up to 10 A/cm² or more.
One of the advantages of a thoriated tungsten cathode over an oxide cathode is the ability to operate the plate at higher voltages. Oxide cathodes are susceptible to deterioration caused by ion bombardment. To achieve reasonable life, plate voltage must be limited. A thoriated tungsten cathode is more tolerant of ion bombardment, and so higher plate voltages can be safely applied.

The end of useful life for a thoriated tungsten tube occurs when most of the carbon has evaporated or has combined with residual gas, depleting the carbide surface layer. Theoretically, a 3% increase in filament voltage will result in a 20 K increase in cathode temperature, a 20% increase in peak emission, and a 50% decrease in tube life because of carbon loss. This cycle works in reverse, too. For a small decrease in temperature and peak emission, the life of the carbide layer—and hence, the tube—may be increased.

**Tungsten-Impregnated Cathode**

The tungsten-impregnated cathode typically operates at 900–1175°C and provides the highest average emission density of the three types of cathodes discussed in this chapter (500 mA/cm²–8 A/cm²). Peak power performance ranges up to 12 A/cm².

Tungsten as an element is better able than other emitters to withstand bombardment by high-energy positive ions without having its emission impaired. These positive ions are always present in small numbers in vacuum tubes as a result of ionization by collision with the residual gas.

**Cathode Construction**

Power tube filaments can be assembled in several different configurations. Figure 5.23 shows a spiral-type filament, and Fig. 5.24 shows a bar-type design. The spiral filament is used extensively in low-power tubes. As the size of the tube increases, mechanical considerations dictate a bar-type filament with spring loading to compensate for thermal expansion. A mesh filament can be used for both small and large tubes. It is

![FIGURE 5.23 Spiral-type tungsten filament. (Source: Varian/Eimac.)](image-url)
more rugged than other designs and less subject to damage from shock and vibration. The rigidity of a cylindrical mesh cathode is determined by the following parameters:

- The diameter of the cathode
- The number, thickness, and length of the wires forming the cathode
- The fraction of welded to total wire crossings

A mesh cathode is shown in Fig. 5.25.

Some power tubes are designed as a series of electron gun structures arranged in a cylinder around a centerline. This construction allows large amounts of plate current to flow and to be controlled with a minimum of grid interception. With reduced grid interception, less power is dissipated in the grid structures. In the case of the control grid, less driving power is required for the tube.

In certain applications, the construction of the filament assembly can have an effect on the performance of the tube itself, and the performance of the RF system as a whole. For example, filaments built in a basket-weave mesh arrangement usually offer lower distortion in critical high-level AM modulation circuits.

**Velocity of Emission**

The electrons emitted from a hot cathode depart with a velocity that represents the difference between the kinetic energy possessed by the electron just before emission and the energy that must be given up to escape. Because the energy of different electrons within the emitter is not the same, the velocity of emission will vary as well, ranging from zero to a maximum value determined by the type and style of emitter.

**Grid Structures**

The type of grid used for a power tube is determined principally by the power level and operating frequency required. For most medium-power tubes (5–25 kW dissipation) welded wire construction is common.
At higher power levels, laser-cut pyrolytic graphite grids may be found. The grid structures of a power tube must maintain their shape and spacing at elevated temperatures. They must also withstand shock and vibration.

**Wire Grids**

Conventional wire grids are prepared by operators that wind the assemblies using special mandrels (forms) that include the required outline of the finished grid. The operators spot weld the wires at intersecting points, shown in Fig. 5.26. Most grids of this type are made with tungsten or molybdenum, which exhibit stable physical properties at elevated temperatures. On a strength basis, pure molybdenum is generally considered the most suitable of all refractory metals at temperatures between 1600 and 3000°F. The thermal conductivity of molybdenum is more than three times that of iron and almost half that of copper.

Grids for higher power tubes are typically built using a bar-cage type of construction. A number of vertical supports are fastened to a metal ring at the top and to a base cone at the bottom. The lower end of the assembly is bonded to a contact ring. The construction of the ring, metal base cone, and cylindrical metal base give the assembly low lead inductance and low RF resistance.

The external loading of a grid during operation and the proximity of the grid to the hot cathode impose severe demands on both the mechanical stability of the structure and the physical characteristics of its surface. The grid absorbs a high proportion of the heat radiated by the cathode. It also intercepts the
electron beam, converting part of its kinetic energy into heat. Further, high-frequency capacitive currents flowing in the grid results in additional heat.

The end result is that grids are forced to work at temperatures as high as 1500°C. Their primary and secondary emission must, however, be low. To prevent grid emission, high electron affinity must be ensured throughout the life of the tube, even though it is impossible to prevent material evaporated from the cathode from contaminating the grid surface.

In tubes with oxide cathodes, grids made of tungsten or molybdenum wire are coated with gold to reduce primary emission caused by deposition. The maximum safe operating temperature for gold plating, however, is limited (about 550°C). Special coatings have, therefore, been developed for high-temperature applications that are effective in reducing grid emission. In tubes with thoriated tungsten cathodes, grids made of tungsten or molybdenum are coated with proprietary compounds to reduce primary emission.

Primary grid emission is usually low in a thoriated tungsten cathode device. In the case of an oxide cathode, however, free barium can evaporate from the cathode coating material and find its way to the control and screen grids. The rate of evaporation is a function of cathode temperature and time. A grid contaminated with barium will become another emitting surface. The hotter the grid, the greater the emissions.

Pyrolytic Grid

Pyrolytic grids are a high-performance alternative to wire or bar grid assemblies. Used primarily for high-power devices, pyrolytic grids are formed by laser-cutting a graphite cup of the proper dimensions. The computer-controlled laser cuts numerous holes in the cup to simulate a conventional-style grid. Figure 5.27 shows a typical pyrolytic-type grid before and after laser processing.

Pyrolytic (or oriented) graphite is a form of crystallized carbon produced by the decomposition of a hydrocarbon gas at high temperatures in a controlled environment. A layer of pyrolytic graphite is deposited on a special form. The thickness of the layer is proportional to the time that deposition is allowed to continue. The structure and mechanical properties of the deposited graphite depend on the imposed conditions.

Pyrolytic grids are ideal vacuum-tube elements because they do not expand like metal. Their small coefficient of expansion prevents movement of the grids inside the tube at elevated temperatures. This preserves the desired electrical characteristics of the device. Because tighter tolerances can be maintained, pyrolytic grids can be spaced more closely than conventional wire grids. Additional benefits include

- The grid is a single structure having no weld points.
- The grid has a thermal conductivity in two of the three planes nearly that of copper.
- The grid can operate at high temperatures with low vapor pressure.

![FIGURE 5.27](image) Pyrolytic graphite grid: (a) before laser processing, (b) completed assembly. (Source: Varian/Eimac.)
**Grid Physical Structure**

The control, screen, and suppressor grids are cylindrical and concentric. Each is slightly larger than the previous grid, as viewed from the cathode. Each is fastened to a metal base cone, the lower end of which is bonded to a contact ring. Figure 5.28 shows the construction of a typical screen grid assembly. Figure 5.29 illustrates a cutaway view of a tetrode power tube.

The shape of the control grid and its spacing from the cathode defines, in large part, the operating characteristics of the tube. For best performance, the grid must be essentially transparent to the electron path from the cathode to the plate. In a tetrode, the control and screen grids must be precisely aligned to minimize grid current. For pentode tubes, the previous two conventions apply, in addition to the requirement for precise alignment and minimum beam interception for the suppressor grid.
Secondary Emission Considerations

The relationship of the properties of secondary electrons to the grid structures and other elements of a vacuum tube must be carefully considered in any design. As the power capability of a tube increases, the physical size of the elements also increase. This raises the potential for secondary emission from the control, screen, and suppressor grids. Secondary emission can occur regardless of the type of cathode used. The yield of secondary electrons may be reduced through the application of surface treatments.

In a tetrode, the screen is operated at a relatively low potential, necessary to accelerate the electrons emitted from the cathode. Not all electrons pass through the screen on their way to the plate. Some are intercepted by the screen grid. As the electrons strike the screen, other low-energy electrons are emitted. If these electrons have a stronger attraction to the screen, they will fall back to that element. If, however, they pass into the region between the screen and the plate, the much higher anode potential will attract them. The result is electron flow from screen to plate.

Because of the physical construction of a tetrode, the control grid will have virtually no control over screen-to-plate current flow as a result of secondary electrons. During a portion of the operating cycle of the device, it is possible that more electrons will leave the screen grid than will arrive. The result will be a reverse electron flow on the screen element, a condition common to high-power tetrodes. A low-impedance path for reverse electron flow must be provided.

Tube manufacturers typically specify the recommended value of bleeding current from the screen power supply to counteract the emission current. Two common approaches are illustrated in Fig. 5.30 and Fig. 5.31. If the screen power supply impedance is excessively high in the direction of reverse electron flow, the screen voltage will attempt to rise to the plate voltage. Note the emphasis on low impedance in the reverse electron flow direction. Most regulated power supplies are low impedance in the forward electron flow direction only. If the supply is not well bled, the reverse electrons will try to flow from anode to cathode in the regulator series pass element. As the screen voltage rises, the secondary and plate currents will increase, and the tube will enter a runaway condition.

As shown in Fig. 5.30, the addition of a 12.5-kΩ resistor from screen to ground provides a path for screen grid emission (20 mA for the circuit shown). In the circuit of Fig. 5.31, plate current flows through the screen power supply, swamping the screen supply. The screen power supply must, obviously, carry the normal screen and plate currents. This scheme is used extensively in circuits where the screen is operated at DC ground potential. The plate-to-cathode voltage is then the sum of the $E_L$ and $E_{c2}$ power supplies.
The suppressor grid of a pentode reduces the effects of secondary emission in the device. This attribute, thus, reduces the requirement to provide a reverse electron flow path for the screen grid power supply. The screen current requirement for a pentode may, however, be somewhat higher than for a tetrode of the same general characteristics.

The designer must also consider the impedance of the control grid circuit. Primary grid emission can result in operational problems if the grid circuit impedance is excessively high. Primary grid emission, in the case of an oxide cathode tube, will increase with tube life.

The size and power of gridded tubes dictate certain characteristics of electrical potential. As this geometry increases in electrical terms, significant secondary emission from the control grid can occur. Control grid secondary emission can exist whether the cathode is a thoriated tungsten or oxide emitter, and can occur in a triode, tetrode, or pentode. A typical curve of grid current as a function of grid voltage for a high-power thoriated tungsten filament tetrode is shown in Fig. 5.32. As shown in the figure, grid current decreases and eventually takes a reverse direction as the grid voltage increases. This reduction and reversal of grid current can be explained by the normal secondary emission characteristics of the metals used in the grid structure. The secondary emission characteristics of common metals are presented in curve form in Fig. 5.33. The ratio of secondary-to-primary electron current is given as a function of primary electron potential. An examination of the chart will show the region between 200 and 600 V to be rather critical as far as secondary emission is concerned. Any power grid tube that normally operates with 200–600 V on the grid can exhibit the negative resistance characteristic of decreasing grid current with increasing grid voltage when another electrode, such as the anode in a triode or the screen grid in a tetrode, is at a sufficiently high potential to attract the emitted electrons from the control grid. A driver stage that works into such a nonlinear load normally must be designed in such a manner as to tolerate this condition. One technique involves swamping the driver so that changes in load resulting from secondary grid emission is a small percentage of the total load the driver works into.

Plate Assembly

The plate assembly of a power tube is typically a collection of many smaller parts that are machined and assembled to tight specifications. Copper is generally used to construct the anode. It is an excellent
material for this purpose because copper has the highest electrical conductivity of any metal except pure silver. Further, copper is easily fabricated and ideally suited to cold-forming operations such as deep drawing, bending, and stamping.

The anode and cooling fins (in the case of an air-cooled device) begin as flat sheets of copper. They are stamped by the tube manufacturer into the necessary sizes and shapes. After all of the parts have been machined, the anode and cooling fins are stacked in their proper positions, clamped, and brazed into one piece in a brazing furnace.

The plate of a power tube resembles a copper cup with the upper-half of a plate contact ring welded to the mouth and cooling fins silver soldered or welded to the outside of the assembly. The lower-half of the anode contact ring is bonded to a base ceramic spacer. At the time of assembly, the two halves of the ring are welded together to form a complete unit, as shown in Fig. 5.34.

In most power tubes, the anode is a part of the envelope, and, because the outer surface is external to the vacuum, it can be cooled directly. Glass envelopes were used in older power tubes. Most have been replaced, however, with devices that use ceramic as the envelope material.

**Ceramic Elements**

Ceramics are an integral part of modern power vacuum tubes. Three types of ceramics are in common usage in the production of vacuum devices:

- *Aluminum oxide*-based ceramics
- *Beryllium oxide*- (BeO-) based ceramics
- *Aluminum nitride*- (AIN-) based ceramics

**Aluminum Oxide Ceramics**

Aluminum oxide-based ceramic insulators are a common construction material for a wide variety of electrical components, including vacuum tubes. Aluminum oxide is 20 times higher in thermal conductivity than most oxides. The flexure strength of commercial high-alumina ceramics is 2–4 times greater than that of most oxide ceramics. There are drawbacks, however, to the use of alumina ceramics, including:

- Relatively high thermal expansion (approximately 7 ppm/°C), compared to other ceramic materials, such as BeO
- Moderately high dielectric constant (approximately 10)

Aluminas are fabricated from aluminum oxide powders with various percentages of sintering promoters, which produce the so-called glassy phase. The later additives reduce the densification temperatures to
between 1500 and 1600°C. Based on the final application, the powders may be pressed, extruded, or prepared in slurries for slip casting or tape casting. The surface finish of aluminas is typically 3–25 µm/in as a result of normal processing. For very smooth finishes (2 µm/in) the surfaces can be lapped or polished.

Alumina ceramics for vacuum tube applications are rarely used apart from being bonded to metals. The means by which this is accomplished frequently dictate the processing technique. Metallization of aluminas is usually accomplished by either high-temperature firing or low-temperature thick-film processing. Fired, shaped aluminas are usually postmetallized by reﬁring the formed article after coating with a slurry of molybdenum and manganese powder or tungsten metal powder. Based on the purity of the alumina, glass powder may be added to the metal powder. The mechanism of metallization requires an internal glassy phase or the added glassy phase for proper bonding. This is accomplished by firing in slightly reducing and moist atmospheres at temperatures above 1500°C. The resulting metallization is usually plated electrochemically with nickel or copper, or both.

Thick-film processing of alumina ceramics has traditionally been performed in oxidizing atmospheres at moderate temperatures (800–1000°C). Precious metals are used, such as gold, silver, platinum, palladium, and their combinations. A glassy phase is usually incorporated in the thick-film paste.

Each approach to metallization has advantages and disadvantages. The advantage of the high-temperature metallization schemes with molybdenum or tungsten is their moderate cost. The disadvantage is the high resistivity of the resulting ﬁlms. These relative merits are reversed for thick-film materials. Another advantage of thick-film metallization is that the process is often applied by the device fabricator, not the ceramic vendor. This can allow for greater efficiency in design modiﬁcation.

**Beryllium Oxide Ceramics**

Beryllium oxide-based ceramics are in many ways superior to alumina-based ceramics. The major drawback is the toxicity of BeO. Beryllium and its compounds are a group of materials that are potentially hazardous and must be handled properly. With the necessary safeguards, BeO has been used successfully in many tube designs.

Beryllium oxide materials are particularly attractive for use in power vacuum tubes because of their electrical, physical, and chemical properties. The thermal conductivity of BeO is approximately 10 times higher than that of alumina-based materials. Figure 5.35 compares the thermal conductivity of BeO to that of alumina and some alternative materials. As the chart illustrates, BeO has a lower dielectric constant and coefﬁcient of thermal expansion than alumina. It is, however, also slightly lower in strength.

Beryllia materials are fabricated in much the same way as alumina compounds, although the toxic properties of the powders mandate that they be processed in laboratories equipped to handle them safely.

---

Simple cutting, drilling, and postmetallization are also handled by vendors with special equipment. Reliable thick-film systems may be applied to BeO substrates. Such coatings usually necessitate less elaborate safety precautions and are often applied by the device fabricator.

**Other Ceramics**

Aluminum nitride- (AIN-) based ceramics have been developed as an alternative to the toxicity concerns of BeO-based materials. As shown in Fig. 5.35, the thermal conductivity of AIN is comparable to that of BeO but deteriorates less with temperature. The dielectric constant of AIN is comparable to that of alumina (a drawback) but its thermal expansion is low (4 ppm/°C).

Other ceramic materials that find some use in vacuum devices include silicon carbide- (SiC-) based substances and boron nitride ceramics.

**Tube Construction**

Each type of power grid tube is unique insofar as its operating characteristics is concerned. The basic physical construction, however, is common to most devices. A vacuum tube is built in essentially two parts:

- The base, which includes the filament and supporting stem, control grid, screen grid, and suppressor grid (if used)
- The anode, which includes the heat-dissipating fins made in various machining steps

The base subassembly is welded using a tungsten-inert gas (TIG) process in an oxygen-free atmosphere (a process sometimes referred to as Heliarc welding) to produce a finished base unit.

The ceramic elements used in a vacuum tube are critical elements of the device. Assembled in sections, each element builds on the previous one to form the base of the tube. The ceramic-to-metal seals are created using a material that is painted onto the ceramic and then heated in a brazing oven. After preparation in a high-temperature oven, the painted area provides a metallic structure that is molecularly bonded to the ceramic, and provides a surface suitable for brazing.

This process requires temperature sequences that dictate completion of the highest temperature stages first. As the assembly takes form, lower oven temperatures are used so that completed bonds will not be damaged.

Despite all of the advantages that make ceramics one of the best insulators for a tube envelope, their brittleness is a potential cause of failure. The smallest cracks in the ceramic, not in themselves damaging, can cause the ceramic to break when mechanically stressed by temperature changes.

After the base assembly has been matched with the anode, the completed tube is brazed into a single unit. The device then goes through a bake-out procedure. Baking stations are used to evacuate the tube and bake out any oxygen or other gases from the copper parts of the assembly. Although oxygen-free copper is used in tube construction, some residual oxygen exists in the metal and must be driven out for long component life. A typical vacuum reading of 10⁻⁸ torr is specified for most power tubes (formerly expressed as 10⁻⁸ mm of mercury). For comparison, this is the degree of vacuum in outer space about 200 mi above the Earth.

A vacuum offers excellent electrical insulation characteristics. This property is essential for reliable operation of a vacuum tube, the elements of which typically operate at high potentials with respect to each other and to the surrounding environment. An electrode containing absorbed gases, however, will exhibit reduced breakdown voltage because the gas will form on the electrode surface, increasing the surface gas pressure and lowering the breakdown voltage in the vicinity of the gas pocket.

To maintain a high vacuum during the life of the component, power tubes contain a getter device. The name comes from the function of the element: to get or trap and hold gases that may evolve inside the tube. Materials used for getters include zirconium, cerium, barium, and titanium.

The operation of a vacuum tube is an evolving chemical process. End of life in a vacuum tube is generally caused by loss of emission.
Connection Points

The high-power levels and frequencies at which vacuum tubes operate place stringent demands on the connectors used to tie the outside world to the inside elements. Figure 5.36 shows a cutaway view of the base of a tetrode. Tubes are designed to be mounted vertically on their electrical connectors. The connectors provide a broad contact surface and mechanical support for the device.

The cathode and grids are mounted on ring-shaped Kovar™ bases, which also serve as contact rings for the external connections. Kovar is an iron-nickel-cobalt alloy whose coefficient of thermal expansion is comparable with that of aluminum oxide ceramic. The different diameters of the various contact rings allow them to be grouped coaxially. The concentric tube/connector design provides for operation at high frequencies. Conductivity is improved by silver plating.

Tube Sockets

Any one tube design may have several possible socket configurations, depending on the frequency of operation. If the tube terminals are large cylindrical surfaces, the contacting portions of the socket consist of either spring collets or an assembly of spring finger stock. Usually, these multiple-contacting surfaces are made of beryllium copper to preserve spring tension at the high temperatures present at the tube terminals. The fingers are silver plated to reduce RF resistance.

If the connecting fingers of a power-tube socket fail to provide adequate contact with the tube element rings, a concentration of RF currents will result. Depending on the extent of this concentration, damage may result to the socket. After a connecting finger loses its spring action, the heating effect is aggravated and tube damage is possible.

A series of specialized power tubes is available with no sockets at all. Intended for cathode-driven service, the grid assembly is formed into a flange that is bolted to the chassis. The filament leads are connected via studs on the bottom of the tube. Such a configuration completely eliminates the tube socket. This type of device is useful for low-frequency applications, such as induction heating.

5.2.3 High Frequency Operating Limits

Like most active devices, performance of a given vacuum tube deteriorates as the operating frequency is increased beyond its designed limit. Electron transit time is a significant factor in the upper frequency limitation of electron tubes. A finite time is taken by electrons to traverse the space from the cathode, through the grid, and on to the plate. As the operating frequency increases, a point is reached at which the electron transit time effects become significant. This point depends on the accelerating voltages at the grid and anode, and their respective spacings. Tubes with reduced spacing in the grid-to-cathode region exhibit reduced transit time effects.
There is also a power limitation that is interrelated with the high-frequency limit of a device. As the operating frequency is increased, closer spacing and smaller-sized electrodes must be used. This reduces the power handling capability of the tube. Figure 5.37 illustrates the relationship.

Gridded tubes at all power levels for frequencies up to about 1 GHz are invariably cylindrical in form. At higher frequencies, planar construction is almost universal. As the operating frequency is increased beyond design limits, output power and efficiency both decrease. Figure 5.38 illustrates the relationship.

Transit Time Effects

When class C, class B, or similar operations are carried out at frequencies sufficiently high that the transit time of the electrons is not a negligible fraction of the waveform cycle, the following complications are observed in grid-based vacuum tubes:

- **Back heating** of the cathode
- Loading of the control grid circuit as a result of energy transferred to electrons that do not necessarily reach the grid to produce a DC grid current

**FIGURE 5.37** Continuous-wave output power capability of a gridded vacuum tube.

**FIGURE 5.38** Performance of a class C amplifier as the operating frequency is increased beyond the design limits of the vacuum tube.
Debunching of plate current pulses

Phase differences between the plate current and the exciting voltage applied to the control grid

Back heating of the cathode occurs when the transit time in the grid-cathode space is sufficiently great to cause an appreciable number of electrons to be in transit at the instant the plate current pulse would be cut off in the case of a low-frequency operation. A considerable fraction of the electrons thus trapped in the interelectrode space are returned to the cathode by the negative field existing in the grid-cathode space during the cutoff portion of the cycle. These returning electrons act to heat the cathode. At very high frequencies this back heating is sufficient to supply a considerable fraction of the total cathode heating required for normal operation. Back heating may reduce the life of the cathode as a result of electron bombardment of the emitting surface. It also causes the required filament current to depend on the conditions of operation within the tube.

Energy absorbed by the control grid as a result of input loading is transferred directly to the electron stream in the tube. Part of this stream is then used by the electrons in producing back heating of the cathode. The remainder affects the velocity of the electrons as they arrive at the anode of the tube. This portion of the energy is not necessarily all wasted. In fact, a considerable percentage of it may, under favorable conditions, appear as useful output in the tube. To the extent that this is the case, the energy supplied by the exciting voltage to the electron stream is simply transferred directly from the control grid to the output circuits of the tube without amplification.

An examination of the total time required by electrons to travel from the cathode to the anode in a triode, tetrode, or pentode operated as a class C amplifier reveals that the resulting transit times for electrons at the beginning, middle, and end of the current pulse will differ as the operating frequency is increased. In general, electrons traversing the distance during the first segment of the pulse will have the shortest transit time, whereas those near the middle and end of the pulse will have the longest transit times, as illustrated in Fig. 5.39. The first electrons in the pulse have a short transit time because they approach the plate before the plate potential is at its minimum value. Electrons near the middle of the pulse approach the plate with the instantaneous plate potential at or near minimum and, consequently, travel less rapidly in the grid-plate space. Finally, those electrons that left the cathode late in the current pulse (those just able to escape being trapped in the control grid-cathode space and returned toward the cathode) will be slowed as they approach the grid, and so have a large transit time. The net effect is to cause the pulse of plate current to be longer than it would in operation at a low frequency. This causes the efficiency of the amplifier to drop at high frequencies, because a longer plate current pulse increases plate losses.

FIGURE 5.39 Transit time effects in a class C amplifier: (a) control grid voltage, (b) electron position as a function of time (triode case), (c) electron position as a function of time (tetrode case), (d) plate current (triode case). (Source: Adapted from Terman, F.E. 1947. Radio Engineering, p. 405. McGraw-Hill, New York.)
Defining Terms

**Back heating:** The effect on the cathode of a power vacuum tube when the transit time in the grid-cathode space is sufficiently great to cause, under certain conditions, an appreciable number of electrons to be trapped in the interelectrode space and returned to the cathode, thereby heating the cathode.

**Bake-out:** A process during tube manufacture in which the device is heated to a high temperature while the envelope is evacuated.

**Collet:** A circular spring fingerstock connection element for a power vacuum tube.

**Electron transit time:** The finite time taken by electrons to traverse the space from the cathode, through the grid(s), and on to the plate of a vacuum tube.

**Pyrolytic grid:** A grid structure made of pyrolytic (oriented) graphite that is laser-cut to the proper geometry for use in a power vacuum tube.

**Thoriated tungsten:** A form of atomic-film emitter commonly used in power tubes. Through proper processing, the hybrid material is made to provide high-emissions capabilities and long life.

**Vapor cooling:** A cooling technique for power vacuum tubes utilizing the conversion of hot water to steam as a means of safely conducting heat from the device and to a heat sink.

References


Brush. Ceramic products. Brush Wellman, Cleveland, OH.


Powers, M.B. *Potential Beryllium Exposure While Processing Beryllia Ceramics for Electronics Applications.* Brush Wellman, Cleveland, OH.

5.3 Neutralization Techniques

Jerry C. Whitaker

5.3.1 Introduction

An RF power amplifier must be properly neutralized to provide acceptable performance in most applications. The means to accomplish this end can vary considerably from one design to another. An RF amplifier is neutralized when two conditions are met:

- The interelectrode capacitance between the input and output circuits is canceled.
- The inductance of the screen grid and cathode assemblies (in a tetrode) is canceled.

Cancellation of these common forms of coupling between the input and output circuits of vacuum tube amplifiers prevents self-oscillation and the generation of spurious products.

5.3.2 Circuit Analysis

Figure 5.40 illustrates the primary elements that effect neutralization of a vacuum tube RF amplifier operating in the VHF band. (Many of the following principles also apply to lower frequencies.) The feedback elements include the residual grid-to-plate capacitance $C_{gp}$, plate-to-screen capacitance $C_{ps}$, and screen grid lead inductance $L_s$. The RF energy developed in the plate circuit $E_p$ causes a current $I$ to flow through the plate-to-screen capacitance and the screen lead inductance. The current through the screen inductance develops a voltage $-E$ with a polarity opposite that of the plate voltage $E_p$. The $-E$ potential is often used as a method of neutralizing tetrode and pentode tubes operating in the VHF band.

Figure 5.41 graphically illustrates the electrical properties at work. The circuit elements of the previous figure have been arranged so that the height above or below the zero potential line represents magnitude.

---

2 Portions of this chapter were adapted from: Varian. 1982. Care and Feeding of Power Grid Tubes. Varian/Eimac, San Carlos, CA. Used with permission.
and polarity of the RF voltage for each part of the circuit with respect to ground (zero). For the purposes of this illustration, assume that all of the circuit elements involved are pure reactances. The voltages represented by each, therefore, are either in phase or out of phase and can be treated as positive or negative with respect to each other.

The voltages plotted in the figure represent those generated as a result of the RF output circuit voltage $E_p$. No attempt is made to illustrate the typical driving current on the grid of the tube. The plate $P$ has a high positive potential above the zero line, established at the ground point. Keep in mind that the distance above the baseline represents increasing positive potential. The effect of the out-of-phase screen potential developed as a result of inductance $L_s$ is shown, resulting in the generation of $-E$.

As depicted, the figure constitutes a perfectly neutralized circuit. The grid potential rests at the zero baseline. The grid operates at filament potential insofar as any action of the output circuit on the input circuit is concerned.

The total RF voltage between plate and screen is made up of the plate potential and screen lead inductance voltage, $-E$. This total voltage is applied across a divider circuit that consists of the grid-to-plate capacitance and grid-to-screen capacitance ($C_{gp}$ and $C_{gs}$). When this potential divider is properly matched for the values of plate RF voltage ($E_p$) and screen lead inductance voltage ($-E$), the control grid will exhibit zero voltage difference with respect to the filament as a result of $E_p$.

**Circuit Design**

There are a variety of methods that may be used to neutralize a vacuum tube amplifier. Generally speaking, a grounded-grid, cathode-driven triode can be operated into the VHF band without external neutralization components. The grounded-grid element is sufficient to prevent spurious oscillations. Tetrode amplifiers generally will operate through the MF band without neutralization. However, as the gain of the stage increases, the need to cancel feedback voltages caused by tube interelectrode capacitances and external connection inductances becomes more important. At VHF frequencies and above, it is generally necessary to provide some form of stage neutralization.

**Below VHF**

For operation at frequencies below the VHF band, neutralization for a tetrode typically employs a capacitance bridge circuit to balance out the RF feedback caused by residual plate-to-grid capacitance. This method assumes that the screen is well bypassed to ground, providing the expected screening action inside the tube.

Neutralization of low-power push-pull tetrode or pentode tubes can be accomplished with cross neutralization of the devices, as shown in Fig. 5.42. Small-value neutralization capacitors are used. In some cases, neutralization can be accomplished with a simple wire connected to each side of the grid circuit and brought through the chassis deck. Each wire is positioned to look at the plate of the tube on the opposite-half of the circuit. Typically, the wire (or a short rod) is spaced a short distance from the plate of each tube. Fine adjustment is accomplished by moving the conductor in or out from its respective tube.

A similar method of neutralization can be used for a cathode-driven symmetrical stage, as shown in Fig. 5.43. Note that the neutralization capacitors $C_n$ are connected from the cathode of one tube to the plate of the opposite tube. The neutralizing capacitors have a value equal to the internal cathode-to-plate capacitance of the PA tubes.
In the case of a single-ended amplifier, neutralization can be accomplished using either a push-pull output or push-pull input circuit. Figure 5.44 shows a basic push-pull grid neutralization scheme that provides the out-of-phase voltage necessary for proper neutralization. It is usually simpler to create a push-pull network in the grid circuit than the plate because of the lower voltages present. The neutralizing capacitor \( C_n \) is small and may consist of a simple feed-through wire (described previously). Padding capacitor \( C_p \) is often added to maintain the balance of the input circuit while tuning. \( C_p \) is generally equal in size to the input capacitance of the tube.

Single-ended tetrode and pentode stages can be neutralized using the method shown in Fig. 5.45. The input resonant circuit is placed above ground by a small amount because of the addition of capacitor \( C_{in} \).
The voltage to ground that develops across $C_{in}$ upon the application of RF drive is out of phase with the grid voltage, and is fed back to the plate through $C_n$ to provide neutralization. In such a design, $C_n$ is considerably larger in value than the grid-to-plate interelectrode capacitance.

The single-ended grid neutralization circuit is redrawn in Fig. 5.46 to show the capacitance bridge that makes the design work. Balance is obtained when the following condition is met

$$\frac{C_n}{C_{in}} = \frac{C_{gp}}{C_{gf}}$$

where

- $C_n$ = neutralization capacitance
- $C_{in}$ = input circuit bypass capacitor
- $C_{gp}$ = grid-to-plate interelectrode capacitance
- $C_{gf}$ = total input capacitance, including tube and stray capacitance

A single-ended amplifier also can be neutralized by taking the plate circuit slightly above ground and using the tube capacitances as part of the neutralizing bridge. This circuit differs from the usual RF amplifier design in that the plate bypass capacitor is returned to the screen side of the screen bypass capacitor, as shown in Fig. 5.47. The size of screen bypass capacitor $C_s$ and the amount of stray capacitances in $C_p$ are chosen to balance the voltages induced in the input by the internal tube capacitances, grid-to-plate $C_{gp}$ and screen-to-grid $C_{sg}$. This circuit is redrawn in Fig. 5.48 in the usual bridge form. Balance is obtained when the following condition is met

$$\frac{C_p}{C_s} = \frac{C_{gp}}{C_{sg}}$$

In usual tetrode and pentode structures, the capacitance from screen-to-grid is approximately half the published tube input capacitance. The tube input capacitance is primarily the sum of the grid-to-screen capacitance and the grid-to-cathode capacitance.
It should be noted that in the examples given, it is assumed that the frequency of operation is low enough so that inductances in the socket and connecting leads can be ignored. This is basically true in MF applications and below. At higher bands, however, the effects of stray inductances must be considered, especially in single-ended tetrode and pentode stages.

**VHF and Above**
Neutralization of power-grid tubes operating at VHF frequencies provides special challenges and opportunities to the design engineer. At VHF frequencies and above, significant RF voltages can develop in the residual inductance of the screen, grid, and cathode elements. If managed properly, these inductances can be used to accomplish neutralization in a simple, straightforward manner.

At VHF and above, neutralization is required to make the tube input and output circuits independent of each other with respect to reactive currents. Isolation is necessary to ensure independent tuning of the input and output. If variations in the output voltage of the stage produce variations of phase angle of the input impedance, phase modulation will result.

As noted previously, a circuit exhibiting independence between the input and output circuits is only half of the equation required for proper operation at RF frequencies. The effects of incidental inductance of the control grid must also be canceled for complete stability. This condition is required because the suppression of coupling by capacitive currents between the input and output circuits is not, by itself, sufficient to negate the effects of the output signal on the cathode-to-grid circuit. Both conditions, input and output circuit independence and compensation for control grid lead inductance, must be met for complete stage stability at VHF and above.

Figure 5.49 shows a PA stage employing stray inductance of the screen grid to achieve neutralization. In this grounded-screen application, the screen is tied to the cavity deck using six short connecting straps. Two additional adjustable ground straps are set to achieve neutralization.

Triode amplifiers operating in a grounded-grid configuration offer an interesting alternative to the more common grounded-cathode system. When the control grid is operated at ground potential, it serves to
shield capacitive currents from the output to the input circuit. Typically, provisions for neutralization are not required until the point at which grid lead inductance becomes significant.

**Grounded-Grid Amplifier Neutralization**

Grounded-grid amplifiers offer an attractive alternative to the more common grid-driven circuit. The control grid is operated at RF ground and serves as a shield to capacitive currents from the output to the input circuit. Generally, neutralization is not required until the control grid lead inductive reactance becomes significant. The feedback from the output to the input circuit is no longer the result of plate-to-filament capacitance. The physical size of the tube and the operating frequency determine when neutralization is required.

Two methods of neutralization are commonly used with grounded-grid amplifiers. In the first technique, the grids of a push-pull amplifier are connected to a point having zero impedance to ground, and a bridge of neutralizing capacitances is used that is equal to the plate-filament capacitances of the tubes.

The second method of neutralization requires an inductance between the grid and ground or between the grids of a push-pull amplifier of a value that will compensate for the coupling between input and output circuits resulting from the internal capacitances of the tubes.

---

**FIGURE 5.49** A grounded-screen PA stage neutralized through the use of stray inductance between the screen connector and the cavity deck: (a) circuit diagram, (b) mechanical implementation.
Behavior of these two circuits is quite different. They may be considered as special forms of the more general case in which the neutralizing capacitors have values differing from the internal capacitances of the tubes, and in which an appropriate reactance is connected between the grids. Under these conditions, the value of neutralizing capacitance permits continuous variation of power amplification, stability, and negative feedback.

The purpose of neutralization is to make the input and output circuits independent of each other with respect to reactive currents. The input current must be independent of the output voltage, and vice versa. This condition is necessary to permit independent tuning of the input and output circuits, so that variations in output voltage do not produce variations of phase angle of the input impedance, resulting in phase modulation.

The condition of independence between input and output circuits, which may be called the neutralized condition, does not necessarily imply stability. This is because the suppression of coupling by capacitive currents between input and output circuits is not sufficient to remove the effect of the output voltage on the cathode-to-grid voltage. A second condition, distinct from neutralization, must be met for complete stability: the effect of the control grid lead inductance must be canceled.

**Grid Impedance**

In the special case of a grounded-grid amplifier having a grid impedance and the reactive currents neutralized, the following equations apply (see Fig. 5.50):

\[
C_n = C_{fp} - \frac{C_{fg}}{\mu} = \frac{1}{j\omega C_{fg} + C_{gp}(1 + \mu)}
\]

If in solving the equation for \( C_n \) the sign is negative, this indicates that in-phase neutralization is required. Conversely, if the sign of \( C_n \) is positive, then out-of-phase neutralization is needed. A negative value of \( Z_g \) indicates capacitive reactance required, and a positive value indicates that inductive reactance is to be used.

**Application Example**

If the grids of a push-pull cathode-driven amplifier are not at ground potential because the inductance of the leads is not negligible, coupling may exist between the input and output circuits through the plate-grid capacitances, cathode-grid capacitances, and grid-to-grid inductance. One method of reducing this coupling is to insert between the grids a series tuned circuit that has zero reactance at the operating frequency. This technique is illustrated in Fig. 5.51. This neutralization scheme is useful only for the case where no grid current flows. If grid current flows, a grid resistance will appear in parallel with the grid-to-filament capacitance. If the resistance is small in comparison to the reactance of this grid-to-filament capacitance, phase modulation will result.

Another important property of the preceding neutralization scheme is that power amplification is a function of the neutralizing capacitance, whereas the independence of cathode and plate circuits from...
FIGURE 5.51 Neutralization by cross-connected capacitors of a symmetrical cathode-excited amplifier with compensation of lead inductance.

the viewpoint of reactive currents may be obtained with any value of neutralizing capacitance. If the neutralizing capacitance is less than the plate-to-filament capacitance of the tube, the stage will operate with low-excitation power and high-power amplification. If the neutralizing capacitance is greater than the plate-to-filament capacitance, the power amplification would be quite low, but the total output power possibly would be increased.

5.3.3 Self-Neutralizing Frequency

The voltage dividing action between the plate-to-grid capacitance \( C_{pg} \) and the grid-to-screen capacitance \( C_{gs} \) will not change with changes in operating frequency. The voltage division between the plate and screen, and screen and ground caused by the charging current \( I_s \) will however, vary significantly with frequency. There will be a particular frequency, therefore, where this potential dividing circuit will effectively place the grid at filament potential insofar as the plate is concerned. This point is known as the self-neutralizing frequency illustrated in Fig. 5.52.

At the self-neutralizing frequency, the tetrode or pentode is inherently neutralized by the circuit elements within the device itself, and external screen inductance to ground. When a device is operated below its self-neutralizing frequency, the normal cross-neutralization circuits apply. When the operating frequency is above the self-neutralizing frequency, the voltage developed in the screen lead inductance is too large to give the proper voltage division between the internal capacitances of the device. One approach to neutralization in this case involves adjusting the inductive reactance of the screen lead to ground so as to lower the total reactance. In the example shown in Fig. 5.53, this is accomplished with a series variable capacitor.

Another approach is shown in Fig. 5.54, in which the potential divider network made up of the tube capacitance is changed. In the example, additional plate-to-grid capacitance is added external to the tube. The external capacitance \( C_{ext} \) can take the form of a small wire or rod positioned adjacent to the plate of the tube. This approach is similar to the one described previously for conventional neutralization, except
that in this case the neutralizing probe is connected to the grid of the tube, rather than to an opposite polarity in the circuit.

If the RF power amplifier is operating above the self-neutralizing frequency of the tube and must be tuned over a range of frequencies, it is probably easier to use the screen series tuning capacitor method and make this control available to the operator. If operation is desired over a range of frequencies including the self-neutralizing frequency of the tube, this circuit is also desirable because the incidental lead inductance in the variable capacitor lowers the self-neutralizing frequency of the circuit so that the neutralizing series capacitor can be made to operate over the total desired frequency range. If this range is too great, switching of neutralizing circuits will be required. A small 50–100 pF variable capacitor in the screen lead has often been found to be satisfactory. Another method of changing the self-neutralizing frequency of a tetrode or pentode can be fashioned from the general bypassing arrangement of the screen and filament shown in Fig. 5.55. The screen lead is bypassed with minimum inductance to the filament terminal of the tube. Some inductance is introduced in the common filament and screen ground leads. The grid is shown below the zero voltage or chassis potential, indicating that the voltage developed in the screen lead inductance to chassis is excessive. If the filament is tapped on this inductance, a point can be found where the voltage difference between the grid and filament is zero, as far as the components of plate voltage are concerned. This arrangement will be found to self-neutralize at a higher frequency than if the filament and screen were separately bypassed to

**FIGURE 5.53** Components of the output voltage of a tetrode when neutralized by added series screen-lead capacitance.

**FIGURE 5.54** Components of the output voltage of a tetrode when neutralized by added external grid-to-plate capacitance.

**FIGURE 5.55** Components of the output voltage of a tetrode neutralized by adding inductance common to the screen and cathode return.
the chassis. Thus, by increasing the self-neutralizing frequency of the tube and screen bypass arrangement, the tendency of the VHF parasitic to occur is reduced.

If the frequency of the VHF parasitic is reduced by increasing the inductance of the plate lead (presuming this is the principle frequency-defining circuit), the circuit can be made to approach the self-neutralizing frequency of the tube and, therefore, suppress the parasitic.

5.3.4 Neutralization Adjustment

Most neutralization circuits must be adjusted for operation at a given frequency. The exact procedure followed to make these adjustments vary from one circuit to the next. The following generalizations, however, apply to most systems.

The first step in the process of neutralization is to break the DC connections to the plate voltage and screen voltage supplies, leaving the RF circuits intact. If the DC current path is not broken, some current can flow in either one of these circuits even though the voltages are zero. The presence of this current causes the amplifier to work in the normal manner, generating RF power in the plate circuit. It will then be incorrect to adjust for zero power in the plate circuit. Sufficient RF grid drive must be applied to provide some grid current or to cause a sensitive RF meter coupled to the plate to give an indication of feedthrough power. When the plate circuit is tuned through resonance, the grid current will dip when the circuit is out of neutralization or the RF meter will peak. The neutralization adjustments are made until the indication is minimum.

Another powerful tool for roughly neutralizing an RF amplifier is to feed the power output from a signal generator into the grid circuit. A sensitive RF detector is inserted between the output connector and the load. Neutralization can then be adjusted for minimum feedthrough. This technique is useful in working with prototype equipment. Actual qualitative measurements can be made. If the insertion loss of the amplifier is less than the expected gain, oscillation will occur. Circuit modifications can be made until the isolation is sufficient to warrant a test with high voltages applied. The advantages of this cold system test include

- No components are subjected to unusual stress if the amplifier is unstable.
- Circuit adjustments can be made safely because no high voltages are present.

For final trimming of the neutralization adjustment, the stage should be returned to operating condition at reduced power (similar to that used when testing for parasitic oscillations), or under the final loaded operating conditions. At the higher frequencies and in the VHF region, it will be found that a small additional trimming adjustment of the neutralization circuit is usually required. When the plate circuit is tuned through resonance, minimum plate current and maximum control grid current should occur simultaneously. In the case of the tetrode and pentode, the DC screen current should be maximum at the same time.

These neutralizing procedures apply not only to the HF radio frequencies, but also in the VHF or UHF regions. In the latter cases, the neutralizing circuit is different and the conventional cross-neutralization schemes may not apply.

Defining Terms

**Interelectrode capacitance**: The effective capacitances between the internal elements of a power vacuum tube. In a triode these typically include the cathode-to-grid capacitance, grid-to-plate capacitance, and cathode-to-plate capacitance.

**Neutralization**: The condition of stability and independence between the input and output circuits of a power grid tube stage.

**Self-neutralizing frequency**: The frequency of operation of a power vacuum tube stage where the potential dividing circuit of the inherent device capacitances will effectively place the grid at filament potential insofar as the plate is concerned.
References

Varian. 1982. *Care and Feeding of Power Grid Tubes*. Varian/Eimac, Laboratory Staff, San Carlos, CA.

Further Information

Specific information on neutralization considerations for power vacuum tubes can be obtained from the manufacturers of those devices. More general application information can be found in the following publications:


5.4 Amplifier Systems

Jerry C. Whitaker

5.4.1 Introduction

Any number of configurations may be used to generate RF signals using vacuum tubes. Circuit design is dictated primarily by the operating frequency, output power, type of modulation, duty cycle, and available power supply. Tube circuits can be divided generally by their operating class and type of modulation employed. The angle of plate current flow determines the class of operation:

- Class A = 360° conduction angle
- Class B = 180° conduction angle
- Class C = conduction angle less than 180°
- Class AB = conduction angle between 180° and 360°

The class of operation has nothing to do with whether the amplifier is grid driven or cathode driven. A cathode-driven amplifier, for example, can be operated in any desired class. The class of operation is only a function of the plate current conduction angle. The efficiency of an amplifier is also a function of the plate current conduction angle.

The efficiency of conversion of DC to RF power is one of the most important characteristics of a vacuum tube amplifier circuit. The DC power that is not converted into useful output energy is, for the most part, converted to heat. This heat represents wasted power; the result of low efficiency is increased operating cost for energy. Low efficiency also compounds itself. This wasted power must be dissipated, requiring increased cooling capacity. The efficiency of the amplifier must, therefore, be carefully considered, consistent with the other requirements of the system. Figure 5.56 shows the theoretical efficiency attainable with a tuned or resistive load assuming that peak AC plate voltage is equal to the plate supply voltage.

---

Class A Amplifier
A class A amplifier is used in applications requiring low harmonic distortion output power. A class A amplifier can be operated with low intermodulation distortion in linear RF amplifier service. Typical plate efficiency for a class A amplifier is about 30%. Power gain is high because of the low drive power required. Gains as high as 30 dB are typical.

Class B and AB Amplifiers
A class AB power amplifier is capable of generating more power—using the same tube—than the class A amplifier, but more intermodulation distortion will also be generated. A class B RF linear amplifier will generate still more intermodulation distortion, but is acceptable in certain applications. The plate efficiency is typically 66%, and stage gain is about 20–25 dB.

Class C Amplifiers
A class C power amplifier is used where large amounts of RF energy need to be generated with high efficiency. Class C RF amplifiers must be used in conjunction with tuned circuits or cavities, which restore the amplified waveform through the flywheel effect.

The grounded cathode class C amplifier is the building block of RF technology. It is the simplest method of amplifying CW, pulsed, and FM signals. The basic configuration is shown in Fig. 5.57. Tuned input and output circuits are used for impedance matching and to resonate the stage at the desired operating frequency. The cathode is bypassed to ground using low-value capacitors. Bias is applied to the grid as shown. The bias power supply may be eliminated if a self-bias configuration is used. The typical operating efficiency of a class C stage ranges from 65 to 85%.

Figure 5.58 illustrates the application of a zero-bias triode in a grounded-grid arrangement. Because the grid operates at RF ground potential, this circuit offers stable performance without the need for neutralization (at medium frequency (MF) and below). The input signal is coupled to the cathode through a matching network. The output of the triode feeds a pi network through a blocking capacitor.

5.4.2 Principles of RF Power Amplification
In an RF power amplifier, a varying voltage is applied to the control grid (or cathode in the case of a grounded-grid circuit) from a driver stage whose output is usually one of the following:

- Carrier frequency signal only
- Modulation (intelligence) signal only
- Modulated carrier signal
Simultaneous with the varying control grid signal, the plate voltage will vary in a similar manner, resulting from the action of the amplified current flowing in the plate circuit. In RF applications with resonant circuits, these voltage changes are smooth sine wave variations, 180° out of phase with the input. The relationship is illustrated in Fig. 5.59. Note how these variations center about the DC plate voltage and the DC control grid bias. In Fig. 5.60 the variations have been indicated next to the plate voltage and grid voltage scales of a typical constant current curve. At some instant in time, shown as $t$ on the time scales, the grid voltage has a value denoted $e_g$ on the grid voltage sinewave.

Any point on the operating line (when drawn on constant current curves as illustrated in Fig. 5.60) indicates the instantaneous values of plate current, screen current, and grid current that must flow when
these particular values of grid and plate voltage are applied to the tube. Thus, by plotting the values of plate and grid current as a function of time $t$, a curve of instantaneous values of plate and grid current can be produced. Such plots are shown in Fig. 5.61.

By analyzing the plate and grid current values it is possible to predict with accuracy the effect on the plate circuit of a change at the grid. It follows that if a properly loaded resonant circuit is connected to

---

**FIGURE 5.59** Variation of plate voltage as a function of grid voltage.

---

**FIGURE 5.60** Relationship between grid voltage and plate voltage plotted on a constant-current curve.
the plate, a certain amount of RF power will be delivered to that circuit. If the resonant circuit is tuned to the fundamental frequency (the same frequency as the RF grid voltage), the power delivered will be that of the fundamental or principal RF component of plate current. If the circuit is tuned to a harmonic of the grid voltage frequency, the power delivered will be the result of a harmonic component of the plate current.

**Drive Power Requirements**

The technical data sheet for a given tube type provides the approximate drive power required for various operating modes. As the frequency of operation increases and the physical size of the tube structure becomes large with respect to this frequency, the drive power requirements will also increase.

The drive power requirements of a typical grounded-cathode amplifier consist of six major elements:

- The power consumed by the bias source, given by
  \[ P_1 = I_{c1} \times E_{c1} \]

- The power dissipated in the grid as a result of rectified grid current
  \[ P_2 = I_{c1} \times e_{cmp} \]

- The power consumed in the tuned grid circuit
  \[ P_3 = I_{c,\text{rms}}^2 \times R_{rf} \]

- The power loss as a result of transit time effects
  \[ P_4 = \left( \frac{e_{c,\text{rms}}}{R_t} \right)^2 \]

  \( R_t \) is that part of the resistive component of the tube input impedance resulting from transit time effects, and is given by

  \[ R_t = \frac{1}{K g_m f^2 T^2} \]

- The power consumed in that part of the resistive component of the input impedance resulting from cathode lead inductance
  \[ P_5 = \frac{e_{c}}{R_{s}} \]

- Input resistance resulting from the inductance of the cathode leads is found from the following
  \[ R_s = \frac{1}{\omega^2 g_m L_k C_{gk}} \]

- The power dissipated in the tube envelope because of dielectric loss
  \[ P_6 = 1.41 f E_1^2 v \]
where

\[ I_{c1} = \text{DC grid current} \]
\[ E_{c1} = \text{DC grid voltage} \]
\[ E_{\text{cmp}} = \text{maximum positive grid voltage} \]
\[ I_{c,\text{rms}} = \text{rms value of RF grid current} \]
\[ R_{\text{RF}} = \text{RF resistance of grid circuit} \]
\[ \varepsilon_{c,\text{rms}} = \text{rms value of RF grid voltage} \]
\[ R_t = \text{resistance resulting from transit time loading} \]
\[ K = \text{a constant (function of the tube geometry)} \]
\[ g_m = \text{transconductance} \]
\[ f = \text{frequency, Hz} \]
\[ T = \text{transit time, cathode-to-grid} \]
\[ R_s = \text{cathode lead inductance input resistance loading} \]
\[ \omega = 2\pi f \]
\[ L_k = \text{cathode lead inductance, H} \]
\[ C_{gk} = \text{grid-to-cathode capacitance, F} \]
\[ E_1 = \text{voltage gradient, kV/in, rms} \]
\[ \nu = \text{loss factor of dielectric materials} \]

The total drive power \( P_t \) is, then, equal to

\[ P_t = P_1 + P_2 + P_3 + P_4 + P_5 + P_6 \]

Particular attention must be given to grid dissipation when a tube is operated in the VHF and UHF regions. The total driving power required for a given output may be greater than the grid dissipation capability of the device.

**Operational Considerations for VHF and UHF**

When operating a tube in the VHF and UHF regions, several techniques may be applied to minimize the driving power without appreciably affecting plate conversion efficiency. These techniques include

- Use the minimum DC control bias. Frequently, it is advisable to bring the bias down to approximately cutoff.
- Maintain a high value of DC screen voltage, even though it appears to increase the fraction of the cycle during which plate current flows.
- Use the minimum RF excitation voltage necessary to obtain the desired plate circuit performance, even though the DC grid current is considerably lower than would be expected at lower frequencies.
- Keep the cathode lead inductance to the output and input circuits as low as possible. This can be accomplished by (1) using short and wide straps, (2) using two separate return paths for the input and output circuits, or (3) proper choice of cathode bypass capacitor(s).

These techniques do not necessarily decrease the plate efficiency significantly when the circuit is operated at VHF and UHF. The steps should be tried experimentally to determine whether the plate circuit efficiency is appreciably affected. It is usually acceptable—and even preferable—to sacrifice some plate efficiency for improved tube life when operating at VHF and UHF.

Optimum power output at these frequencies is obtained when the loading is greater than would be used at lower frequencies. Fortunately, the same condition reduces driving power and screen current (for the tetrode and pentode cases), and improves tube life expectancy in the process.

**Mechanical and Electrical Considerations**

To maintain proper isolation of the output and input circuits, careful consideration must be given to the location of the component parts of the amplifier. All elements of the grid or input circuit and any earlier
stages must be kept out of the plate circuit compartment. Similarly, plate circuit elements must be kept out of the input compartment. It should be noted, however, that in the case of the tetrode and pentode, the screen lead of the tube and connections via the socket are common to both the output and input resonant circuits. Because of the plate-to-screen capacitance of a tetrode or pentode, the RF plate voltage (developed in the output circuit) causes an RF current to flow out the screen lead to the chassis. In the case of a push-pull stage, this current may flow from the screen terminal of one tube to the screen terminal of the other tube. Similarly, because of the grid-to-screen capacitance of the tube, the RF voltage in the input circuit will cause an RF current to flow in this same screen lead to the chassis, or to the opposite tube of the push-pull circuit.

The inductance of the lead common to both the output and input circuits has the desirable feature of providing voltage of opposite polarity to neutralize the feedback voltage of the residual plate to control grid capacitance of the tube. It should be noted, however, that the mutual coupling from the screen lead to the input resonant circuit may be a possible source of trouble, depending on the design.

**Power Supply Considerations**

The power supply requirements for a triode are straightforward. The degree of regulation and ripple depends on the requirements of the system. In the case of a linear RF amplifier, it is important to have good plate power supply regulation. Without tight regulation, the plate voltage will drop during the time the plate is conducting current heavily. This drop will cause flat topping and will appear as distortion in the output. In push-pull applications where grid current flows, it is important to keep the grid circuit resistance to a minimum. If this is not done, positive peak clipping will occur.

In the case of the tetrode and pentode, the need for screen voltage introduces some new considerations and provides some new possibilities. Voltage for the screen grid of a low-power tetrode or pentode can readily be taken from the power supply used for the plate of the tube. In this case, a series resistor, or potential dividing resistor, is chosen so that with the intended screen current flowing the voltage drop through the resistor is adequate to give the desired screen voltage. The potential dividing resistor is the preferred technique for those tubes with significant secondary screen emission.

It is possible to take the screen voltage from a separate low-voltage supply. A combination scheme may also be employed, where a dropping resistor is used in conjunction with a low-voltage or intermediate-voltage supply. Frequently a combination of series resistor and voltage source can be chosen so that the rated screen dissipation will not be exceeded regardless of variations in screen current. With a fixed screen supply, there are advantages in using an appreciable amount of fixed grid bias so as to provide protection against the loss of excitation, or for cases where the driver stage is being keyed.

If the screen voltage is taken through a dropping resistor from the plate supply, there is usually little point in using a fixed grid bias because an unreasonable amount of bias would be required to protect the tube if the excitation failed. Under operating conditions with normal screen voltage, the cutoff bias is low (screen voltage divided by the screen \( \mu \)). When a stage loses excitation and runs statically, the screen current falls to nearly zero. (See the static curves of the tube in question.) If the screen voltage is obtained through a simple dropping resistor from the plate supply, the screen voltage will then rise close to full plate voltage. Because the cutoff bias required is proportional to screen voltage, the grid bias required will be much greater than the amount of bias desired under normal operating conditions. When a screen dropping resistor is used, most of the bias is normally supplied through a grid resistor and other means are used for tube protection.

The power output from a tetrode or pentode is sensitive to screen voltage. For this reason, any application requiring a high degree of linearity through the amplifier requires a well-regulated screen power supply. A screen dropping resistor from the plate supply is not recommended in such applications.

The suppressor grid power supply requirements are similar to the control grid power supply. The suppressor grid intercepts little current and, therefore, a low power supply may be used. Any variation in suppressor voltage as a result of ripple or lack of regulation will appear at the output of the amplifier because of suppressor grid modulation of the plate current.
Parasitic Oscillations

Most self-oscillations in RF power amplifiers using gridded tubes have been found to fall in the following three classes:

- Oscillation at VHF from about 40 to 200 MHz, regardless of the operating frequency of the amplifier.
- Self-oscillation on the fundamental frequency of the amplifier.
- Oscillation at a low radio frequency below the normal frequency of the amplifier.

Low-frequency oscillation in an amplifier usually involves the RF chokes, especially when chokes are used in both the output and input circuits. Oscillation near the fundamental frequency involves the normal resonant circuits, and brings up the question of neutralization. When a parasitic self-oscillation is found on a very high frequency, the interconnecting leads of the tube, tuning capacitor, and bypass capacitors are typically involved.

VHF oscillation occurs commonly in amplifiers where the radio frequency circuits consist of coils and capacitors, as opposed to cavity designs. As illustrated in Fig. 5.62, the tube capacitances effectively form a tuned-plate tuned-grid oscillator.

The frequency of a VHF parasitic is typically well above the self-neutralizing frequency of the tube. However, if the self-neutralizing frequency of the device can be increased and the frequency of the parasitic lowered, complete suppression of the parasitic may result, or its suppression by resistor-inductor parasitic suppressors may be made easier.

It is possible to predict with the use of a grid dip wave meter the parasitic frequency to be expected in a given circuit. The circuit should be complete and with no voltages on the tube. Couple the meter to the plate or screen lead and determine the resonant frequency.

Elimination of the VHF parasitic oscillation may be accomplished using one of the following techniques:

- Place a small coil and resistor combination in the plate lead between the plate of the tube and the tank circuit, as shown in Fig. 5.63(a). The resistor-coil combination is usually made up of a noninductive resistor of about 25–100 Ω, shunted by three or four turns approximately 1/2-in diameter and frequently wound around the resistor. In some cases it may be necessary to use such a suppressor in both the plate and grid leads. The resistor-coil combination operates on the principle that resistor loads the VHF circuit but is shunted by the coil for the lower fundamental frequency. In the process of adjusting the resistor-coil combination, it is often found that the resistor runs hot. This heat is usually caused by the dissipation of fundamental power in the resistor, and is an indication of too many turns in the suppressor coil. Just enough turns should be used to suppress the parasitic and no more. Once the parasitic has been suppressed there will be no parasitic voltage or current present. Therefore, there is no parasitic power to be dissipated.

- Use small parasitic chokes in the plate lead as shown in Fig. 5.63(b). The size of the coil will vary considerably depending on the tube and circuit layout. A coil of 4–10 turns measuring approximately 1/2-in diameter is typical. The presence of the choke in the frequency determining part of the circuit

![FIGURE 5.62 Inter electrode capacitances supporting VHF parasitic oscillation in an HF RF amplifier.](image-url)
FIGURE 5.63 Placement of parasitic suppressors to eliminate VHF parasitic oscillations in an HF RF amplifier: (a) resistor-coil combination, (b) parasitic choke.

... lowers the frequency of a possible VHF parasitic so that it falls near the self-neutralizing frequency of the tube and bypass leads. In addition to varying the size of the suppressor choke, the amount of inductance common to the screen and filament in the filament grounding strap may be a factor. This parameter can be varied simultaneously with the suppressor choke.

Of the two methods outlined for suppressing VHF parasitic oscillations, the first is probably the simpler to use and has been widely employed.

**Dynatron Oscillation**

Another form of commonly encountered self-oscillation is known as dynatron oscillation. Dynatron oscillation is caused when any electrode in a vacuum tube has negative resistance. At times there may be more electrons leaving the screen grid than are arriving. If the screen voltage is allowed to increase under these conditions, even more electrons will leave the grid. This phenomenon implies a negative resistance characteristic. If there is high alternating current impedance in the circuit from the screen grid through the screen grid power supply, and from the plate power supply to the plate, dynatron oscillation may be sustained.

Dynatron oscillation typically occurs in the region of 1–20 Hz. This low-frequency oscillation is usually accompanied by another oscillation in the 1–2 kHz region. Suppression of these oscillations can be accomplished by placing a large bypass capacitor (1000 µF) across the output of the screen grid power supply. The circuit supporting the oscillation can also be detuned by a large inductor. Increasing the circuit losses at the frequency of oscillation is also effective.

**Harmonic Energy**

It is generally not appreciated that the pulse of grid current contains energy on harmonic frequencies and that control of these harmonic energies may be important. The ability of the tetrode and pentode to isolate the output circuit from the input circuit over a wide range of frequencies is important in avoiding feed through of harmonic voltages from the grid circuit. Properly designed tetrode and pentode amplifiers provide for complete shielding in the amplifier layout so that coupling external to the tube is prevented.

In RF amplifiers operating either on the fundamental or a desired harmonic frequency, the control of unwanted harmonics is important. The following steps permit reduction of the unwanted harmonic energies present in the output circuit:

- Keep the circuit impedance between the plate and cathode low for the high harmonic frequencies. This requirement may be achieved by having some or all of the tuning capacitance of the resonant circuit close to the tube.
- Completely shield the input and output compartments.
- Use inductive output coupling from the resonant plate circuit and possibly a capacitive or Faraday shield between the coupling coil and the tank coil, or a high-frequency attenuating circuit such as a π or π-L network.
· Use low-pass filters on all supply leads and wires coming into the output and input compartments.
· Use resonant traps for particular frequencies.
· Use a low-pass filter in series with the output transmission line.

**Shielding**
In an RF amplifier, shielding between the input and output circuits must be considered. Triode amplifiers are more tolerant of poor shielding because power gain is relatively low. If the circuit layout is reasonable and no inductive coupling is allowed to exist, a triode amplifier can usually be built without extensive shielding. Even if shielding is not necessary to prevent fundamental frequency oscillation, it will aid in eliminating any tendency toward parasitic oscillation. The higher the gain of an amplifier the more important the shielding.

**Pierced Shields**
Tetrode and pentode amplifiers require comprehensive shielding to prevent input-to-output circuit coupling. It is advisable to use nonmagnetic materials such as copper, aluminum, or brass in the RF fields to provide the shielding. Quite often a shield must have holes through it to allow the passage of cooling air. In the LF and part of the HF range, the presence of small holes will not impair shield effectiveness. As the frequency is increased, however, the RF currents flowing around the hole in one compartment cause fields to pass through the hole into another compartment. Currents are, therefore, induced on the shield in the other compartment. This problem can be eliminated by using holes that have significant length. A piece of pipe with a favorable length-to-diameter ratio as compared to the frequency of operation will act as a *waveguide beyond cutoff* attenuator. If more than one hole is required to pass air, a material resembling a honeycomb may be used. This material is commercially available and provides excellent isolation with a minimum air pressure drop. A section of honeycomb shielding is shown in Fig. 5.64. Some tube sockets incorporate a waveguide beyond cutoff air path. These sockets allow the tube in the amplifier to operate at high gain and up through VHF.

**FIGURE 5.64** A section of honeycomb shielding material used in an RF amplifier.
Compartments
By placing the tube and related circuits in completely enclosed compartments and properly filtering incoming supply wires, it is possible to prevent coupling radio frequency energy out of the circuit by means other than the desired output coupling. Such filtering prevents the coupling of energy that may be radiated or fed back to the input section or earlier stages in the amplifier chain. Energy fed back to the input circuit causes undesirable interaction in tuning and/or self-oscillation. If energy is fed back to earlier stages, significant operational problems may result because of the large-power gain over several stages.

In the design of an RF amplifier, doors or removable panels must typically be used. The requirement for making a good, low-resistance joint at the discontinuity must be met. There are several materials available commercially for this purpose. Finger stock, shown in Fig. 5.65, has been used for many years.

Protection Measures
Power grid tubes are designed to withstand considerable abuse. The maximum ratings for most devices are conservative. For example, the excess anode dissipation resulting from detuning the plate circuit will have no ill effects on most tubes if not applied for periods of time sufficient to overheat the envelope and the seal structure.

Similarly, the control, screen, and suppressor grids will stand some excess dissipation. Typically, the maximum dissipation for each grid indicated on the data sheet should not be exceeded except for time intervals of less than 1 s. The maximum dissipation rating for each grid structure is usually considerably above typical values used for maximum output so that ample operating reserve is provided. The time of duration of overload on a grid structure is necessarily short because of the small heat storage capacity of the grid wires. Furthermore, grid temperatures cannot be measured or seen, and so no warning of accidental overload is apparent.

The type and degree of protection required in an RF amplifier against circuit failure will vary with the type of screen and grid voltage supply. Table 5.3 lists protection criteria for tetrode and pentode devices. The table provides guidelines on the location of a suitable relay that should act to remove the principal supply voltage from the stage or transmitter to prevent damage to the tube.

For designs where screen voltage taken through a dropping resistor from the plate supply, a plate relay provides almost universal protection. For the case of a fixed screen supply, a relay provides protection in most cases. For protection against excessive antenna loading and consequent high plate dissipation, a screen undercurrent relay may also be used in some services.

The plate, screen, and bias voltages may be applied simultaneously to a tetrode. The same holds true for a pentode, plus the application of the suppressor voltage. In a grid-driven amplifier, grid bias and excitation can usually be applied alone to the tube, especially if a grid leak resistor is used. Plate voltage can be applied to the tetrode and pentode before the screen voltage with or without excitation to the control grid. Never apply screen voltage before the plate voltage. The only exception would be when the tube is cut off so that no space current (screen or plate current) will flow, or when the excitation and screen voltage are low. If screen voltage is applied before the plate voltage and screen current can
### TABLE 5.3 Protection Guidelines for Tetrode and Pentode Devices

<table>
<thead>
<tr>
<th>Circuit Failure Type</th>
<th>Fixed Screen Supply</th>
<th>Screen Voltage through Dropping Resistor</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Fixed Grid Bias</td>
<td>Resistor Grid Bias</td>
</tr>
<tr>
<td>Loss of excitation</td>
<td>No protection</td>
<td>Plate current relay</td>
</tr>
<tr>
<td></td>
<td>required</td>
<td>Plate current relay or screen control</td>
</tr>
<tr>
<td>Loss of antenna loading</td>
<td>Screen current relay</td>
<td>Grid current relay</td>
</tr>
<tr>
<td>Excess antenna loading</td>
<td>Screen undercurrent relay</td>
<td>No protection required</td>
</tr>
<tr>
<td>Failure of plate supply</td>
<td>Screen current relay</td>
<td>Grid current relay</td>
</tr>
<tr>
<td>Failure of screen supply</td>
<td>Grid current relay</td>
<td>No protection required</td>
</tr>
<tr>
<td>Failure of grid bias supply</td>
<td>Plate current relay</td>
<td>Does not apply</td>
</tr>
<tr>
<td></td>
<td>Plate current relay or screen current relay</td>
<td>Does not apply</td>
</tr>
</tbody>
</table>

flow, the maximum allowable screen dissipation will almost always be exceeded and tube damage will result.

Table 5.4 lists protection guidelines for a triode. The table covers the grid-driven triode amplifier and the high-$\mu$ (zero-bias) cathode-driven triode amplifier. Drive voltage must never be applied to a zero-bias triode amplifier without plate voltage. The table indicates the recommended location of a suitable relay that should act to remove the principal supply voltage from the stage or transmitter to prevent damage to the tube or transmitter.

### 5.4.3 Cavity Amplifier Systems

Power grid tubes are ideally suited for use as the power generating element in a cavity amplifier. Because of the physical dimensions involved, cavity designs are typically limited to VHF frequencies and above. Lower frequency RF generators utilize discrete $L$ and $C$ devices to create the needed tank circuit. Two types of cavity amplifiers are commonly used: 1/4-wavelength and 1/2-wavelength systems.

In a cavity amplifier, the tube becomes part of a resonant transmission line. The stray interelectrode and distributed capacity and inductance of the tube are used to advantage as part of the resonant line. This resonant line is physically larger than the equivalent lumped constant $LRC$ resonant circuit operating at the same frequency, and this larger physical size aids in solving the challenges of high-power operation, skin effect losses, and high-voltage standoff concerns.

A shorted 1/4-wavelength transmission line has a high, purely resistive input impedance. Electrically, it appears as a parallel resonant circuit, as shown in Fig. 5.66. When the physical length of the line is less than
1/4-wavelength, the impedance will be lower and the line will appear inductive, as illustrated in Fig. 5.67. This inductance is used to resonate with the capacitive reactance in the tube and the surrounding circuit.

**Bandwidth and Efficiency**

Power amplifier bandwidth has a significant effect on modulation performance. Available bandwidth determines the amplitude response, phase response, and group delay response. Performance tradeoffs must often be considered in the design of a cavity amplifier, including bandwidth, gain, and efficiency.

Power amplifier bandwidth is restricted by the equivalent load resistance across the parallel tuned circuits in the stage. Tuned circuits are necessary to cancel the low reactive impedance presented by the relatively high input and output capacitances of the amplifying device. The bandwidth for a single tuned circuit is proportional to the ratio of capacitive reactance $X_c$ to load resistance $R_l$ appearing across the tuned circuit

$$BW \approx \frac{K}{2\pi f_c C R_l} \approx \frac{K X_c}{R_l}$$

where

- $BW = \text{bandwidth between half-power (–3 dB) points}$
- $K = \text{proportionality constant}$
- $R_l = \text{load resistance appearing across tuned circuit}$

**FIGURE 5.66** Shorted 1/4-wavelength line: (a) physical circuit, (b) electrical equivalent.

**FIGURE 5.67** Shorted line less than 1/4-wavelength: (a) physical circuit, (b) electrical equivalent.
\[ C = \text{total capacitance of tuned circuit (includes stray capacitances and output or input capacitances of the tube)} \]

\[ X_c = \text{capacitive reactance of } C \]

\[ f_c = \text{carrier frequency} \]

The RF voltage swing across the tuned circuit also depends on the load resistance. For the same power and efficiency, the bandwidth can be increased if the capacitance is reduced.

The efficiency of a cavity power amplifier (PA) depends primarily on the RF plate voltage swing, the plate current conduction angle, and the cavity efficiency. Cavity efficiency is related to the ratio of loaded to unloaded \( Q \) as follows:

\[ N = 1 - \frac{Q_l}{Q_u} \times 100 \]

where

- \( N \) = cavity efficiency,
- \( Q_l \) = cavity loaded \( Q \)
- \( Q_u \) = cavity unloaded \( Q \)

The loaded \( Q \) is determined by the plate load impedance and output circuit capacitance. Unloaded \( Q \) is determined by the cavity volume and the RF resistivity of the conductors resulting from the skin effect. For best cavity efficiency, the following conditions are desirable:

- High unloaded \( Q \)
- Low loaded \( Q \)

As the loaded \( Q \) increases, the bandwidth decreases. For a given tube output capacitance and power level, loaded \( Q \) decreases with decreasing plate voltage or with increasing plate current. The increase in bandwidth at reduced plate voltage occurs because the load resistance is directly related to the RF voltage swing on the tube anode. For the same power and efficiency, the bandwidth can also be increased if the output capacitance is reduced. Power tube selection and minimization of stray capacitance are important considerations when designing for maximum bandwidth.

**Current Paths**

The operation of a cavity amplifier is an extension of the current paths inside the tube. Two elements must be examined in this discussion:

- The input circulating currents
- The output circulating currents

**Input Circuit**

The grid/cathode assembly resembles a transmission line whose termination is the RF resistance of the electron stream within the tube. Fig. 5.68 shows the current path of an RF generator (the RF driver stage output) feeding a signal into the grid/cathode circuit.

The outer contact ring of the cathode heater assembly makes up the inner conductor of a transmission line formed by the cathode and control grid assemblies. The filament wires are returned down the center of the cathode. For the input circuit to work correctly, the cathode must have a low RF impedance to ground. This cathode bypassing may be accomplished in one of several ways.

Below 30 MHz, the cathode can be grounded to RF voltages by simply bypassing the filament connections with capacitors, as shown in Fig. 5.69(a). Above 30 MHz, this technique does not work well because of the stray inductance of the filament leads. Notice that in Fig. 5.69(b), the filament leads appear as RF chokes, preventing the cathode from being placed at RF ground potential. This causes negative feedback and reduces the efficiency of the input and output circuits.
FIGURE 5.68 A simplified representation of the grid input circuit of a PA tube.

FIGURE 5.69 Three common methods for RF bypassing of the cathode of a tetrode PA tube: (a) grounding the cathode below 30 MHz, (b) grounding the cathode above 30 MHz, (c) grounding the cathode via a 1/2-wave transmission line.
In Fig. 5.69(c) the cathode circuit is configured to simulate a 1/2-wave transmission line. The line is bypassed to ground with large value capacitors 1/2-wavelength from the center of the filament (at the filament voltage feed point). This transmission line RF short is repeated 1/2-wavelength away at the cathode (heater assembly) and effectively places it at ground potential.

Because 1/2-wavelength bypassing is usually bulky at VHF frequencies (and may be expensive), RF generators are often designed using certain values of inductance and capacitance in the filament/cathode circuit to create an artificial transmission line that will simulate a 1/2-wavelength shorted transmission line. As illustrated in the figure, the inductance and capacitance of the filament circuit can resemble an artificial transmission line of 1/2-wavelength, if the values of \( L \) and \( C \) are properly selected.

**Output Circuit**

The plate-to-screen circulating current of the tetrode is shown in Fig. 5.70. For the purposes of example, consider that the output RF current is generated by an imaginary current generator located between the plate and screen grid. The RF current travels along the inside surface of the plate structure (because of the skin effect), through the ceramic at the lower-half of the anode contact ring, across the bottom of the fins, and to the band around the outside of the fins. The RF current then flows through the plate bypass capacitor to the RF tuned circuit and load, and returns to the screen grid.

The return current travels through the screen bypass capacitor (if used) and screen contact ring, up the screen base cone to the screen grid, and back to the imaginary generator.

The screen grid has RF current returning to it, but because of the assembly’s low impedance, the screen grid is effectively at RF ground potential. The RF current generator, therefore, appears to be feeding an open-ended transmission line consisting of the anode (plate) assembly and the screen assembly. The RF voltage developed by the anode is determined by the plate impedance \( (Z_p) \) presented to the anode by the resonant circuit and its load.

When current flows on one conductor of a transmission line cavity circuit, an equal magnitude current flows in the opposite direction on the other conductor. This means that a large value of RF circulating current is flowing in a cavity amplifier outer conductor (the cavity box). All of the outer conductor circulating currents start at and return to the screen grid (in a tetrode-based system). The front or back
access panel (door) of the cavity is part of the outer conductor and large values of circulating current flow into it, through it, and out of it. A mesh contact strap is generally used to electrically connect the access panel to the rest of the cavity.

### 1/4-Wavelength Cavity

The 1/4-wavelength PA cavity is common in transmitting equipment. The design is simple and straightforward. A number of variations can be found in different RF generators, but the underlying theory of operation is the same.

A typical 1/4-wave cavity is shown in Fig. 5.71. The plate of the tube connects directly to the inner section (tube) of the plate-blocking capacitor. The exhaust chimney/inner conductor forms the other plate of the blocking capacitor. The cavity walls form the outer conductor of the 1/4-wave transmission line circuit. The DC-plate voltage is applied to the PA tube by a cable routed inside the exhaust chimney and inner tube conductor. In the design shown in the figure, the screen-contact fingerstock ring mounts on a metal plate that is insulated from the grounded cavity deck by an insulating material. This hardware makes up the screen-blocking capacitor assembly. The DC-screen voltage feeds to the fingerstock ring from underneath the cavity deck through an insulated feedthrough assembly.

A grounded-screen configuration may also be used in this design in which the screen-contact fingerstock ring is connected directly to the grounded cavity deck. The PA cathode then operates at below ground potential (in other words, at a negative voltage), establishing the required screen voltage to the tube.

The cavity design shown in the figure is set up to be slightly shorter than a full 1/4-wavelength at the operating frequency. This makes the load inductive and resonates the tube's output capacity. Thus, the physically foreshortened shorted transmission line is resonated and electrically lengthened to 1/4-wavelength.

Figure 5.72 illustrates the paths taken by the RF-circulating currents in the circuit. RF energy flows from the plate, through the plate-blocking capacitor, along the inside surface of the chimney/inner conductor (because of the skin effect), across the top of the cavity, down the inside surface of the cavity box, across the cavity deck, through the screen-blocking capacitor, over the screen-contact fingerstock and into the screen grid.
Figure 5.72  The RF circulating current paths for the 1/4-wavelength cavity shown in Fig. 5.71.

Figure 5.73 shows a graph of RF current, voltage, and impedance for a 1/4-wavelength coaxial transmission line. It shows that infinite impedance, zero RF current, and maximum RF voltage occur at the feed point. This would not be suitable for a practical PA circuit because arcing would result from the high RF voltage, and poor efficiency would be caused by the mismatch between the tube and the load.

Notice, however, the point on the graph marked at slightly less than 1/4-wavelength. This length yields an impedance of 600–800 Ω and is ideal for the PA-plate circuit. It is necessary, therefore, to physically foreshorten the shorted coaxial transmission-line cavity to provide the correct plate impedance. Shortening the line also is a requirement for resonating the tube’s output capacity, because the capacity shunts the transmission line and electrically lengthens it.

Copyright 2005 by Taylor & Francis Group
Electron Vacuum Devices

1/4-WAVELENGTH DISTANCE ALONG TRANSMISSION LINE

AMPLITUDE CURRENT VOLTAGE IMPEDANCE

PLATE SCREEN GRID

FIGURE 5.74 Graph of the RF current, RF voltage, and RF impedance produced by the physically foreshortened coaxial transmission line cavity.

Figure 5.74 shows a graph of the RF current, voltage, and impedance presented to the plate of the tube as a result of the physically foreshortened line. This plate impedance is now closer to the ideal 600–800 Ω value required by the tube anode.

Tuning the Cavity
Coarse tuning of the cavity is accomplished by adjusting the cavity length. The top of the cavity (the cavity shorting deck) is fastened by screws or clamps and can be raised or lowered to set the length of the cavity for the particular operating frequency. Fine tuning is accomplished by a variable-capacity plate-tuning control built into the cavity. In a typical design, one plate of the tuning capacitor, the stationary plate, is fastened to the inner conductor just above the plate-blocking capacitor. The movable tuning plate is fastened to the cavity box, the outer conductor, and is mechanically linked to the front-panel tuning control. This capacity shunts the inner conductor to the outer conductor and is used to vary the electrical length and resonant frequency of the cavity.

The 1/4-wavelength cavity is inductively coupled to the output port. This coupling is usually on the side opposite the cavity access door. The inductive pickup loop can take several forms. In one design it consists of a half-loop of flat copper bar stock that terminates in the loading capacitor at one end and feeds the output transmission line inner conductor at the other end. This configuration is shown in Fig. 5.75. The inductive pickup ideally would be placed at the maximum current point in the 1/4-wavelength cavity. However, this point is located at the cavity shorting deck and when the deck is moved for coarse tuning, the magnetic coupling will be changed. A compromise in positioning, therefore, must be made. The use of a broad, flat copper bar for the coupling loop adds some capacitive coupling to augment the reduced magnetic coupling.

Adjustment of the loading capacitor couples the 50-Ω transmission-line impedance to the impedance of the cavity. Heavy loading lowers the plate impedance presented to the tube by the cavity. Light loading reflects a higher load impedance to the amplifier plate.

Methods commonly used to increase the operating bandwidth of the cavity include minimizing added tuning capacitance. The ideal case would be to resonate the plate capacitance alone with a perfect inductor. Practical cavities, however, require either the addition of a variable capacitor or a variable inductor using...
sliding contacts for tuning. Other inherent mechanical and electrical compromises include:

- The requirement for a plate DC blocking capacitor.
- The presence of maximum RF current at the grounded end of the simulated transmission line where the conductor may be nonhomogeneous. This can result in accompanying losses in the contact resistance.

**Wideband Cavity**

The cavity systems discussed so far in this section provide adequate bandwidth for many applications. Some uses, however, dictate an operating bandwidth beyond what a conventional cavity amplifier can provide. One method of achieving wider operating bandwidth involves the use of a double-tuned overcoupled amplifier, shown in Fig. 5.76. The system includes four controls to accomplish tuning:

- **Primary tune**, which resonates the plate circuit and tends to tilt the response and slide it up and down the bandpass.
- **Coupling**, which sets the bandwidth of the PA. Increased coupling increases the operating bandwidth and lowers the PA efficiency. When the coupling is adjusted, it can tilt the response and change the center of the bandpass, necessitating readjustment of the plate tune control.
- **Secondary tune**, which resonates the secondary cavity and will tilt the response if so adjusted. The secondary tune control typically can be used to slide the response up and down the bandpass, as the primary tune control.
- **Loading**, which determines the value of ripple in the bandpass response. Adjustment of the loading control usually tilts the response and changes the bandwidth, necessitating readjustment of the secondary tune and coupling controls.

**Output Coupling**

Coupling is the process by which RF energy is transferred from the amplifier cavity to the output transmission line. Wideband cavity systems use coupling to transfer energy from the primary cavity to the secondary cavity. Coupling in tube type power amplifiers usually transforms a high (plate or cavity) impedance to a lower output (transmission line) impedance. Both capacitive (electrostatic) and inductive (magnetic) coupling methods are used in cavity RF amplifiers. In some designs, combinations of the two methods are used.
Inductive Coupling

Inductive (magnetic) coupling employs the principles of transformer action. The efficiency of the coupling depends on three conditions:

- The cross-sectional area under the coupling loop, compared to the cross-sectional area of the cavity (see Fig. 5.77). This effect can be compared to the turns ratio principle of a transformer.
- The orientation of the coupling loop to the axis of the magnetic field. Coupling from the cavity is proportional to the cosine of the angle at which the coupling loop is rotated away from the axis of the magnetic field, as illustrated in Fig. 5.78.
- The amount of magnetic field that the coupling loop intercepts (see Fig. 5.79). The strongest magnetic field will be found at the point of maximum RF current in the cavity. This is the area where maximum inductive coupling is obtained. Greater magnetic field strength also is found closer to the center conductor of the cavity. Coupling, therefore, is inversely proportional to the distance of the coupling loop from the center conductor.

FIGURE 5.77 The use of inductive coupling in a 1/4-wavelength PA stage.
In 1/4-wavelength cavities, the coupling loop generally feeds a 50-Ω transmission line (the load). The loop is in series with the load and has considerable inductance at VHF frequencies. This inductance will reduce the RF current that flows into the load, thus reducing power output. This effect can be overcome by placing a variable capacitor in series with the output coupling loop. The load is connected to one end of the coupling loop and the variable capacitor ties the other end of the loop to ground. The variable capacitor cancels some or all of the loop inductance. It functions as the PA-stage loading control.

Maximum loop current and output power occurs when the loading capacitor cancels all of the inductance of the loading loop. This lowers the plate impedance and results in heavier loading.

Light loading results if the loading capacitance does not cancel all of the loop inductance. The loop inductance that is not canceled causes a decrease in load current and power output, and an increase in plate impedance.

**Capacitive Coupling**

Capacitive (electrostatic) coupling, which physically appears to be straightforward, often baffles the applications engineer because of its unique characteristics. Figure 5.80 shows a cavity amplifier with a capacitive-coupling plate positioned near its center conductor. This coupling plate is connected to the
output load, which can be a transmission line or a secondary cavity (for wideband operation). The parameters that control the amount of capacitive coupling include

- The area of the coupling-capacitor plate (the larger the area, the greater the coupling)
- The distance from the coupling plate to the center conductor (the greater the distance, the lighter the coupling)

Maximum capacitive coupling occurs when the coupling plate is at the maximum voltage point on the cavity center conductor.

To understand the effects of the capacitive coupling, the equivalent circuit of the cavity must be observed. Fig. 5.81 shows the PA tube, cavity (functioning as a parallel resonant circuit), and output section. The plate-blocking capacitor isolates the tube's DC voltage from the cavity. The coupling capacitor and output load are physically in series, but electrically they appear to be in parallel, as shown in Fig. 5.82. The resistive component of the equivalent parallel circuit is increased by the coupling reactance. The equivalent parallel coupling reactance is absorbed into the parallel resonant circuit. This explains the necessity to retune the plate after changing the PA stage coupling (loading). The coupling reactance can be a series capacitor or inductor.

The series-to-parallel transformations are described by the following formulas

\[ R_p = \frac{R_s^2 + X_s^2}{R_s} \]
\[ X_p = \frac{R_s^2 + X_s^2}{X_s} \]
**PA Loading**

Proper loading of a cavity PA stage to the output transmission line is critical to dependable operation and optimum efficiency of the overall system. Light coupling produces light loading and results in a high plate impedance; conversely, heavy coupling results in heavier loading and a lower plate impedance. Maximum output power, coinciding with maximum efficiency and acceptable dissipation, dictates a specific plate impedance for a cavity of given design. This plate impedance is also dependent upon the values of DC plate voltage ($E_p$) and plate current ($I_p$).

Plate impedance dictates the cavity parameters of loaded $Q$, RF circulating current, and bandwidth. The relationships can be expressed as follows:

- Loaded $Q$ is directly proportional to the plate impedance and controls the other two cavity parameters. Loaded $Q = Z_p / X_l$, where $Z_p$ is the cavity plate impedance and $X_l$ is the cavity inductive reactance.
- Circulating current in the cavity is much greater (by a factor of the loaded $Q$) than the RF current supplied by the tube. Circulating current = $Q \times I_p$, where $I_p$ is the RF current supplied to the cavity by the tube.
- The cavity bandwidth is dependent on the loaded $Q$ and operating frequency. Bandwidth is equal to $F_r / Q$, where $F_r$ is the cavity resonant frequency.

Heavy loading lowers the PA plate impedance and cavity $Q$. A lower $Q$ reduces the cavity RF circulating currents. In some cavities, high circulating currents can cause cavity heating and premature failure of the plate or screen blocking capacitors. The effects of lower plate impedance, a byproduct of heavy loading, are higher RF and DC plate currents and reduced RF plate voltage. The instantaneous plate voltage is the result of the RF plate voltage added to the DC plate voltage. The reduced swing of plate voltage causes less positive DC screen current to flow. Positive screen current flows only when the plate voltage swings close to or below the value of the positive screen grid.

Light loading raises the plate impedance and cavity $Q$. A higher $Q$ will increase the cavity circulating currents, raising the possibility of component overheating and failure. The effects of higher plate impedance...
are reduced RF and DC plate current and increased RF and DC plate voltage excursions. The higher cavity RF or peak DC voltages may cause arcing in the cavity.

There is one value of plate impedance that will yield optimum output power, efficiency, dissipation, and dependable operation. It is dictated by cavity design and the values of the various DC and RF voltages and currents supplied to the stage.

Depending on the cavity design, light loading may seem deceptively attractive. The DC plate voltage is constant (set by the power supply), and the lower DC plate current resulting from light loading reduces the tube's overall DC input power. The RF output power may change with light loading, depending on the plate impedance and cavity design, while efficiency will probably increase or, at worst, remain constant. Caution must be exercised with light loading, however, because of the increased RF voltages and circulating currents that such operation creates. Possible adverse effects include cavity arcing and overheating of cavity components, such as capacitors. The manufacturer's recommendations on PA tube loading should, therefore, be carefully observed.

Although there are many similarities among various cavity designs, each one imposes its own set of operational requirements and limitations. No two cavity systems will tune up in exactly the same fashion. Given proper maintenance, a cavity amplifier will provide years of reliable service.

**Mechanical Design**

Understanding the operation of a cavity amplifier is usually difficult because of the nature of the major component elements. It is often difficult to relate the electrical schematic diagram to the mechanical assembly that exists within the transmitter. Consider the PA cavity schematic diagram shown in Fig. 5.83. The grounded-screen stage is of conventional design. Decoupling of the high-voltage power supply is accomplished by $C_1, C_2, C_3,$ and $L_1$. Capacitor $C_3$ is located inside the PA chimney (cavity inner conductor). The RF sample lines provide two low-power RF outputs for a modulation monitor or other test instrument. Neutralization inductors $L_3$ and $L_4$ consist of adjustable grounding bars on the screen grid ring assembly.

![Figure 5.83 Typical VHF cavity amplifier.](Diagram)
Figure 5.84 shows the electrical equivalent of the PA cavity schematic diagram. The 1/4-wavelength cavity acts as the resonant tank for the PA. Coarse tuning of the cavity is accomplished by adjustment of the shorting plane. Fine tuning is performed by the PA tuning control, which acts as a variable capacitor to bring the cavity into resonance. The PA loading control consists of a variable capacitor that matches the cavity to the load. The assembly made up of $L_2$ and $C_6$ prevents spurious oscillations within the cavity.

FIGURE 5.85 VHF cavity amplifiers: (a) cross-sectional view of a broadband design, (b) cavity designed for television service, (c) FM broadcast cavity amplifier.
Blocking capacitor $C_4$ is constructed of a roll of Kapton™ insulating material sandwiched between two circular sections of aluminum. (Kapton is a registered trademark of Du Pont.) PA plate tuning control $C_5$ consists of an aluminum plate of large surface area that can be moved in or out of the cavity to reach resonance. PA loading control $C_7$ is constructed much the same as the PA tuning assembly, with a large-area paddle feeding the harmonic filter, located external to the cavity. The loading paddle may be moved toward the PA tube or away from it to achieve the required loading. The $L_2$-$C_6$ damper assembly consists of a $50$-$\Omega$ noninductive resistor mounted on the side of the cavity wall. Component $L_2$ is formed by the inductance of the connecting strap between the plate tuning paddle and the resistor. Component $C_6$ is the equivalent stray capacitance between the resistor and the surrounding cavity box.

It can be seen that cavity amplifiers involve as much mechanical engineering as they do electrical engineering. The photographs of Fig. 5.85 show graphically the level of complexity that a cavity amplifier may involve. Figure 5.85(a) depicts a VHF power amplifier (Philips) with broadband input circuitry. Figure 5.85(b) shows a wideband VHF amplifier intended for television service (Varian). Figure 5.85(c) illustrates a VHF cavity amplifier designed for FM broadcast service (Varian).

**Defining Terms**

**Broadband cavity amplifier:** A cavity amplifier made to operating over a broad range of frequencies, typically through the addition of a secondary cavity that is coupled to the primary.

**Cavity amplifier:** A vacuum tube-based amplifying stage in which the physical elements of the output resonating circuit consist of the tube chimney and enclosing box. These and related components in the output circuit typically simulate a $1/4$- or $1/2$-wavelength transmission line.

**Excitation:** The input signal (typically to the grid of a power vacuum tube) that controls the operation of the stage, usually an amplifier or oscillator.

**Flat topping:** A distortion mechanism in an amplifying stage in which the stage is unable to faithfully reproduce the positive peaks of the output signal. Reasons for such problems include poor regulation of the plate voltage supply.

**Loading:** The process of extracting energy from a cavity amplifier through either inductive or capacitive coupling.

**Operating class:** A division of operating modes for an amplifying device. In the case of power vacuum tubes, the operating class is determined by the angle of plate current flow as follows: class $A = 360^\circ$ conduction angle, class $B = 180^\circ$ conduction angle, class $C =$ conduction angle less than $180^\circ$, and class $AB =$ conduction angle between $180$ and $360^\circ$.

**References**


5.5 Image Capture Devices

Steve Epstein

5.5.1 Introduction

The invention of image capture tubes, and cathode ray tube (CRT) technology for display, is at the heart of what eventually became television. Near the end of the 19th century, when early work was being done on these devices, generating artificial images (i.e., test patterns) simply was not practical. Image capture devices allowed real world images to be captured and turned into an electronic signal. Early attempts were mechanical, followed later by electronic vacuum tube devices. As tube technology improved, picture resolution increased to the point where, during the 1960s and 1970s, a variety of tubes were found in professional video cameras throughout the world.

Ultimately tube technology improved to the point where reasonably high-resolution images (>750 TV lines) could be obtained with compact, lightweight devices. These units found their way into the electronic news gathering (ENG) cameras of the mid-1970s. These cameras, along with portable videocassette recorders, changed the face of television news forever. No longer burdened by the requirement of film, television news became more immediate. Stories shot within minutes of newscasts could be edited and aired.

By the 1980s, solid-state charge-coupled devices (CCDs) were finding their way into cameras. Early CCDS, like early tubes, had their problems. Today, however, CCDS offer high resolution, long-life, high sensitivity, and negligible maintenance costs. Tube devices did see a short resurrection during the early design phase of high-definition television, however, increased resolution CCDS became available and high definition development work quickly moved toward the CCD. The majority of the image pickup tubes available in quantity today are of the Plumbicon™ variety.

5.5.2 The Plumbicon

The development of the Plumbicon in 1965 (Philips Components) was an enormous technological achievement. At a time when color television was just getting off the ground, the Plumbicon paved the way for its explosive growth. The Plumbicon and similar tubes, including the Leddicon™ (English Electric Value Company, Ltd. (EEV)), used a lead-oxide photoconductive layer. Lead oxide is a solid-state semiconductor, a porous vaporgrown crystalline layer of lead monoxide.

The basic physical construction of the Plumbicon is shown in Fig. 5.86. The photosensitive PbO layer is deposited on a transparent electrode of SnO₂ on the window of the tube. This conducting layer is the
target electrode and is typically biased at about 40 V positive with respect to the cathode. The remainder of the tube structure is designed to provide a focused, low-energy electron beam for scanning the target.

Figure 5.87 provides a more detailed schematic of the operation of the Plumbicon. The side of the PbO layer facing the gun is reduced to cathode potential by the charge deposited by the scanning beam. Thus, the full target voltage appears across the layer. In the absence of an optical signal (lower-half of figure), there is no mechanism for discharging, and the surface charges to such a potential that further charge deposition, often called beam landing, is prevented. At this point, there is no further current induced in the external circuit. When light is incident as in the upper-half of the figure, photocarriers are generated. Under the influence of the electric field across the layer, these carriers move in the appropriate direction to discharge the stored charge on the surface. As the scanning beam passes across the varied potential of the surface, an additional charge can be deposited on the higher positive potential regions. The amount of charge deposited is equal to that discharged by the light. The attendant current in the external circuit is the photosignal.

In the Plumbicon, the lead oxide layer acts as a $p-i-n$ diode. The $n$-region is formed at the PbO : SnO$_2$ interface (window side). The $p$-region is a thin layer on the electron-gun side of the layer. The bulk of the material behaves as if it were intrinsic, the $i$-layer. Thus, the electron beam landing on the $p$-region charges it negatively with respect to the $n$-region. The $p-i-n$ layer is thus biased in the reverse or low conduction direction, and in the absence of illumination, little or no current flows.

Figure 5.88 shows a simplified diagram of a reverse biased $p-i-n$ diode. Virtually all of the applied voltage $V$ appears across the $i$-region. The dark current of this device is determined by the inverse currents through the $p-i$ and $i-n$ junctions and the thermal generation of carriers in the $i$-region. To ensure that the dark current in the tube due to these sources is sufficiently low at ambient temperatures, a semiconductor of bandgap greater than 1.0 eV is required. For the somewhat elevated temperatures that are likely to be encountered in a camera, a still larger bandgap is needed. The bandgap of the tetragonal lead oxide used in a Plumbicon is about 1.9 eV, and extremely low dark currents result.

When the target structure is illuminated, the amount of current is dependent on photon energy and wavelength. Both blue and green light generate reasonable amounts of current. Plumbicons, however, are not particularly responsive to wavelengths in the red region. To improve the red response photoconductors
are doped, usually with sulfur. These specially prepared tubes, often called extended reds, are then used in the red channel to improve camera performance (Fig. 5.89).

5.5.3 Operating Characteristics

For camera tubes, two important parameters to consider are modulation depth and lag. Modulation depth is determined by the spot diameter of the electron beam and by the thickness of the photoconductive layer. Lag is governed by the process of recharging the photoconductive layer. This is accomplished by the electron beam. The speed of the process is limited by the capacitance of the photoconductive layer and the differential resistance of the electron beam. The photoconductive layer is essentially capacitive in nature; as the layer thickness increases, the capacitance decreases. As the capacitance is reduced, response time for recharging is improved, but due to additional light dispersion, depth of modulation decreases.

Differential beam resistance is determined by the velocity spread of individual electrons in the beam and by the charging current. Less velocity spread gives a lower differential beam resistance and, therefore, faster response. Early Plumbicons used a conventional triode gun, whereas later versions took advantage of a diode gun. The diode gun was able to provide reduced velocity spread and, therefore, offered an improvement in lag characteristics. Lag can be further reduced through the use of a bias light. A bias light is placed at the base of the tube and projects light onto the photoconductive layer. This has the effect of offsetting the zero point of the charging current.

Designing for Tubes

For tubes, factors that influence resolution include the following.

- Lenses: Lenses have their own modulation transfer function (MTF) characteristics that change with aperture of \( f \) setting.
- Wavelength of the light: The color of the light will affect lens MTF, which will influence resolution.

![Figure 5.88](image1.png)

**FIGURE 5.88** Simplified energy band characteristics of a reverse biased \( p-i-n \) diode.

![Figure 5.89](image2.png)

**FIGURE 5.89** Differences in the light sensitivity of (1) standard and (2) extended red Plumbicons.
Electron Vacuum Devices

- Imaging tube properties: The photosensitive layer and the beam spot size will affect maximum resolution.
- Yoke properties: The predominant function of the yoke is to provide the scanning action of the beam and to focus the beam. The focus field, along with the G2, G3, and G4 electrodes, forms an electromagnetic lens. This lens influences the beam shape at the target. Focus field strength in the vicinity of the target strongly influences spot size, as does the voltage ratio of G4 to G3.
- Flare or veiling glare: Flare can reduce resolution by softening edges. Various methods are employed to reduce flare including special coatings on lens surfaces and antihalation pills on the front surface of Plumbicons.
- Signal preamps: Preamp frequency response directly affects output resolution. Preamps should be swept with a device that accurately simulates the tube to verify that high-frequency noise is minimal and that overall response is flat.
- Noise: Noise is the electrical equivalent of attempting to view objects through a snowstorm. Resolution is reduced by the noise. Poor signal-to-noise response results in pictures with apparently poor resolution due to the noise inherent in the image.
- Beam alignment: If beam alignment is not optimized, the beam will not land perpendicular to the center of the target, thereby reducing center resolution.

When designing circuitry for tubes, it is best to consult the manufacturer for data specific to the device to be used. For general purposes, the specifications of a generic 1.5-in tube will be discussed here so that designers will be familiar with the concepts involved. Because a 1.5-in tube is fairly large, voltage and current specifications will generally be higher than what would be required in smaller tubes. Like most tubes, imaging tubes require relatively high voltage and significant amounts of current and, generally, produce heat that must be dissipated. Because circuit requirements and output characteristics can vary with temperature, dissipation must be considered at the design stage.

Imaging tubes generally have seven or eight pins on their base, arranged around the center. Many times a slot is left for a ninth pin, which is used as an alignment reference. Other tubes have a registration mark on the side. This mark is used to align the tube within the deflection yokes. This orients the internal tube structures within the external magnetic fields such that tube performance is maximized. Bias lights are internal on some tubes, others have external units that fit over the base pins and remain in place when the socket is installed. Most bias lights use the heater supply as their power source, adding approximately 100 mA each to the power supply requirements. In existing cameras adapted for operation with a bias light, the added current is not usually a problem and modifications are not usually required. Additional heating is negligible.

Imaging tubes require heaters that normally draw less than 1 A at 6 V. Cathode voltage is generally 0 V; however, most are driven positive during blanking periods. Current requirements are usually less than 250 nA. The target is typically biased at +40 V for Plumbicons and up to +125 V for vidicons. The signal electrode is attached to the target and is generally brought out of the glass envelope near the front of the unit rather than through the base at the rear. Because of the small currents involved, the target ring and its mating connector within the camera head must be kept clean.

Up to four internal grid structures are connected to pins located on the base along with the heater and cathode pins. Grid 1 (G1) is used to blank the beam when necessary, and is typically biased from 0 (beam on) to −100 V (beam off). G2 is used to accelerate the beam and is biased at +300 V or thereabouts. G3 is used to focus the beam and is normally set from +800 to +1000 V. G4 is the mesh assembly located adjacent to the target and is typically biased at +1400 V.

External structures are also required for proper tube operation. These structures fall into two categories, mechanical and electrical. Mechanical structures include the mounting assembly and its associated adjustments. In a typical application, a locking unit holds the tube and its electromagnetic coils rigidly in place. When loosened, the tube and coils can be adjusted within a small range. One adjustment moves the tube back and forth to obtain optimal optical focus, the other rotates the assembly to align the tube with the horizontal and vertical axes of the image. An additional unit, usually a locking ring, is used to hold the tube in place within the deflection coil assembly.
External electrical structures include the deflection coils, external focusing coils, and alignment coils (Fig. 5.90). For the focusing coil, field strength should be of the order of 4–5 mT. Alignment coils are used to correct slight misalignments that result from tube and yoke manufacturing. Typical field strengths required are less than 0.5 mT. Two coils are generally used: one to correct horizontally, the other to correct vertically. Deflection coils are used to deflect the beam in such a manner as to produce a raster. For a typical national television system committee (NTSC) raster, coil current for horizontal deflection will be on the order of 250 mA, with another 50 mA required for vertical deflection.

Other considerations include the ability to beam down and cap the tubes when they are not in use. Tubes have a finite lifetime; eventually cathode emissions fall to a point where the tube becomes unusable. Circuitry that turns off the beam when not in use can extend tube life considerably. A capping mechanism, either in the camera’s filter wheel or a physical lens cap, should be used to prevent highlights from reaching the tube when the camera is not in operation.

Tubes should not be subjected to vibration or shock when oriented vertically with the base in the uppermost position. Under these circumstances, particles can be dislodged within the tube. These loose particles can land on the target, causing spots or blemishes. Large particles landing on the target within the image area can effectively ruin an otherwise good tube. Optical quality glass is used on the surface of the tube and care should be taken to ensure it is kept clean and scratch free.

**Defining Terms**

**Anticomet tail (ACT):** A special type of electron gun designed to handle highlights by increasing beam current with a defocused beam during line retrace.

**Blooming:** An area of the target that is unstable due to insufficient beam current. The area normally appears as a white puddle without definition. Insufficient beam current may be the result of a low beam control setting.

**Dark current:** Current flow from the photosensitive layer in the absence of incident light. This could be compared to leakage current through a reverse-biased diode.

**Diode gun Plumbicon:** A Plumbicon tube with an electron gun that operates with a positive voltage applied to G₁ with respect to the cathode. The diode gun principle provides a finer beam spot size and lower beam temperature. This results in higher resolution and improved lag performance compared to triode gun tubes. The diode gun also provides a much higher current reserve for highlight handling when used in conjunction with a **dynamic beam control** (DBC) circuit.

**Dynamic beam control:** A circuit in a camera designed to instantaneously increase beam current to handle highlights in a scene. This is an alternative to the ACT solution of highlight control.

**Dynamic resolution:** The ability to distinguish fine detail in a moving object in a scene.

**High stability (HS) diode gun Plumbicon:** A diode gun Plumbicon tube, with electrostatic focus and magnetic deflection, which uses a high stability electrode structure evaporated and bonded to the tube envelope.
Image retention: A short-term after-image created by a high-contrast stationary scene within the beaming capabilities of the tube.

Lag: The inability of an imaging tube to respond to instantaneous changes in light. For measurement purposes, lag has two components: rise lag is the response time from dark to light, whereas decay lag is the response time from light to dark. Lag is a very short-term effect, and should not be confused with image retention, image burn, or sticking.

Low output capacitance (LOC) Plumbicon tube: The LOC Plumbicon tube is designed to reduce the capacitance of the target to ground, resulting in an improved signal-to-noise ratio.

MS diode gun Plumbicon: A diode gun Plumbicon tube with magnetic focus and electrostatic deflection.

Modulation depth: The ratio of the peak-to-peak amplitudes of, for example, 0.5–5 MHz (40–400 television lines) signal as measured on a waveform monitor. When measuring modulation depth, consideration of the frequency response of the preamplifiers as well as the MTF of the lens must be taken into account.

Noise: The lower limit of useable dynamic range of the camera. Signal-to-noise is defined as the ratio of the peak value of the maximum signal [100 Institute of Radio Engineers (IRE) units] to the RMS noise voltage when the camera is capped.

Picture quality: Cleanliness of the scanned raster area with respect to spots and blemishes.

Resolution: The ability to distinguish fine detail in a video scene. The resolution of a camera tube is commonly expressed in terms of modulation depth.

Sensitivity: The signal current developed per unit of incident light, measured in microamps per lumen.

Shading: The variation of sensitivity across the target with reference to image center.

Stabilized beam current: The amount of beam current required to stabilize the target when a given amount of light is incident on the target. The beam current is normally set at two times picture white (commonly referred to as one F stop past 100 IRE or one F stop of head room).

Sticking: Very short-term image retention created in a high-contrast moving or stationary scene. Sticking takes on the appearance of lag except that it occurs under normal lighting conditions, whereas lag is generally a low light level phenomenon.

Triode gun: An electron gun that operates with a negative voltage applied to G1 with respect to the cathode.

References

Levitt, R.S. 1968. Operating Characteristics of the Plumbicon.

In addition, the following books are recommended.


Copyright 2005 by Taylor & Francis Group
Further Information

Additional information on the topic of image capture devices is available from the following sources.

*Broadcast Engineering* magazine is a monthly periodical dealing with television and radio technology. The magazine, published in Overland Park, KS, is free to qualified subscribers.

The Society of Motion Picture and Television Engineers (SMPTE) publishes a monthly *Journal*, and holds an annual technical conference in February and a convention in the Fall. The SMPTE is headquartered in White Plains, NY.

The Society of Broadcast Engineers (SBE) holds technical conferences throughout the year. The SBE is located in Indianapolis, IN.

The National Association of Broadcasters (NAB) holds an annual engineering conference and trade show in the Spring. The NAB is headquartered in Washington, D.C.

### 5.6 CRT Devices and Displays

*Jerry C. Whitaker*

#### 5.6.1 Introduction

The cathode ray tube (CRT) is the dominant display technology for a wide range of applications—both consumer and professional.\(^4\) As the requirements for greater resolution and color purity have increased, improvements have also been made in the design and manufacture of CRT devices and signal-driving circuits. Improvements to the basic monochrome and/or color CRT have been pushed within the last 10 years by the explosion of the personal computer industry and the increased resolution demanded by end users. Display size has also been a key element in CRT development. Consumer interest in large-screen home television has been strong within the last decade, and the roll-out of high-definition television has accelerated this trend.

#### 5.6.2 Basic Operating System

The CRT produces visible or ultraviolet radiation by bombardment of a thin layer of phosphor material by an energetic beam of electrons. Nearly all commercial applications involve the use of a sharply focused electron beam directed time sequentially toward relevant locations on the phosphor layer by means of externally controlled electrostatic or electromagnetic fields. In addition, the current in the electron beam can be controlled or modulated in response to an externally applied varying electric signal. A generalized CRT consists of the following elements:

- An electron beam-forming system
- Electron-beam deflecting system (electrostatic or electromagnetic)
- Phosphor screen
- Evacuated envelope

Figure 5.91 shows the basic design of a monochrome CRT. The electron beam is formed in the electron gun, where it is modulated and focused. The beam then travels through the deflection region, where it is directed toward a specific spot or sequence of spots on the phosphor screen. At the phosphor screen the electron beam gives up some of the energy of the electrons in producing light or other radiation, some in generating secondary electrons, and the remainder in producing heat.

---

Classification of CRT Devices

Tubes may be classified in terms of bulb parameters and screen/gun geometry. The principle categories that separate one class of device from another include:

- **Tube size.** Conventionally, tube size is measured as the screen diagonal dimension in rounded inch units. This number is included in tube-type numbers.
- **Neck diameter (OD).** The gun, yoke, neck hardware, and socketing are affected by this dimension (typically given in millimeter). Common neck sizes are 36.5, 29, and 22.5 mm.
- **Deflection angle.** This parameter is calculated from the rated full-screen diagonal and glassware drawings, using the yoke reference plane as an assumed center of deflection. Angles in common use include 90, 100, and 110°. Higher deflection angles enable shorter tubes but entail other tradeoffs.
- **Other characteristics,** including gun type (delta or in-line), screen structure (stripes or dots) for color CRTs, and flat or curved face plates.

The CRT Envelope

The cathode ray tube envelope consists of the faceplate, bulb, funnel, neck, base press, base, faceplate safety panel, shielding, and potting. (Not all CRTs incorporate each of these components.) The faceplate is the most critical component of the envelope because the display on the phosphor must be viewed through it. Most faceplates are pressed in molds from molten glass and trimmed and annealed before further processing. Some specialized CRTs for photographic recording or flying-spot scanning use optical-quality glass faceplates sealed to the bulb section in such a way as to produce minimum distortion.

To minimize the return scattering of ambient light from the white phosphor, many CRT types use a neutral-gray-tinted faceplate. While the display information is attenuated as it makes a single pass through this glass, ambient light will be attenuated both going in and coming out, thus squaring the attenuation ratio and increasing contrast.

Certain specialized CRTs have faceplates made wholly or partially of fiber optics, which may have extraordinary characteristics, such as high ultraviolet transmission. A fiber optic region in the faceplate permits direct-contact exposure of photographic or other sensitive film without the necessity for external lenses or physical space for optical projection.

The bulb section of the CRT is the transition element necessary to enclose the full deflection volume of the electron beam between the deflection region and the phosphor screen on the faceplate. In most
CRTs, this is a roughly cone-shaped molded-glass component. Instead of an ordinary glass bulb, many large CRTs include metal cone sections made of a glass-sealing iron alloy. The metal cones are generally lighter in weight than the corresponding glass sections.

The junction region of the bulb of a CRT with the neck section is critical to the geometry of the device. Tubes made with these separate sections are intended for electromagnetic deflection, and this region is where the deflection yoke is located.

The neck diameter of a CRT depends to a great extent on the type of deflection used and the intended application of the CRT. In general, electrostatic deflection CRTs have large neck diameters, whereas the electromagnetic-deflection devices have small diameters.

**Phosphor and Screen Characteristics**

Many materials, naturally occurring and synthetic, organic, and inorganic, have the ability to give off light. For video applications, the materials of interest are crystalline inorganic solids that are stable under cathode-ray tube fabricating and operating conditions. These materials are generally powders having average particle sizes in the range of 5–15 µm. Figure 5.92 shows some typical particle size distributions. Because of defects and irregularities in the crystal lattice structure, these materials have the ability to absorb incident energy and convert this energy into visible light. This process involves the transfer of energy from the electron beam to electrons in the phosphor crystal. The phosphor electrons are thereby excited or raised to levels higher than the ground state. Light is emitted when the electrons return to the more stable states. Figure 5.93 illustrates these changes in energy levels.

Phosphors are composed of a host crystal that comprises the bulk of the material and one or more activators, which may be present in amounts from parts per million to a few mole percent. Either the host or activator can determine the luminescent properties of a phosphor system. For example, in the zinc sulfide/cadmium sulfide: silver (ZnS:Ag/CdS:Ag) system, the emitted color ranges from blue at zero cadmium through green, to yellow, and into red as the cadmium content is increased.

In the commercial preparation of phosphors, the highly purified host and the required amount of activator are intimately mixed, normally with a flux, such as an alkali or alkaline earth halide or phosphate, which supplies a low-temperature melting phase. The flux controls the particle development and aids in the diffusion of the activator into the lattice. This mixture is then fired at high temperature, 1472–2192°F (800–1200°C) on a prescribed schedule in order to develop the desired physical and luminescent properties.
In some cases the firing is carried on under specific controlled atmospheres. After firing, the resultant cake is broken up, residual soluble materials are removed by washing, and any required coatings are applied.

**Phosphor Screen Types**

A phosphor screen is used to convert electron energy into radiant energy in a CRT display device. The screen is composed of a thin layer of luminescent crystals, phosphors, that emit light when bombarded by electrons. This property is referred to as cathodoluminescence. It occurs when the energy of the electron beam is transferred to electrons in the phosphor crystal. The property of light emission during excitation is termed fluorescence, and that immediately after excitation is removed is termed phosphorescence.

Standardization of phosphor types has been accomplished by registration of the various phosphors with the Joint Electron Device Engineering Council (JEDEC) of the Electronic Industries Association (EIA). Registered phosphors are designated by a number series known as *P numbers*, as shown in Table 5.5.

**Luminescent Properties**

Phosphors are commercially available with cathodoluminescent emission over the entire visible band, including ultraviolet and near-infrared. Table 5.6 lists the important characteristics of the most common phosphors used in video display applications. Absolute phosphor efficiency is measured as the ratio of total absolute energy emitted to the total excitation energy applied. When evaluating or comparing picture

---

**TABLE 5.5** Characteristics of Standard Phosphor Types

<table>
<thead>
<tr>
<th>EIA no.</th>
<th>Typical Application</th>
<th>Composition</th>
<th>Relative Efficiency, %</th>
<th>x^a</th>
<th>y^a</th>
<th>u^a</th>
<th>v^a</th>
<th>Decay</th>
</tr>
</thead>
<tbody>
<tr>
<td>P-4 WW</td>
<td>Black and white television</td>
<td>ZnS:Ag + Zn(_{1-x})Cd(_x)S:Cu, Al</td>
<td>100</td>
<td>0.270</td>
<td>0.300</td>
<td>0.178</td>
<td>0.446</td>
<td>Medium</td>
</tr>
<tr>
<td>P-1 GJ</td>
<td>Projection green</td>
<td>ZnSiO(_2):Mn</td>
<td>130</td>
<td>0.218</td>
<td>0.712</td>
<td>0.079</td>
<td>0.577</td>
<td>Medium</td>
</tr>
<tr>
<td>P-43 GY</td>
<td>Projection green</td>
<td>Gd(_2)O(_2):Tb</td>
<td>155</td>
<td>0.333</td>
<td>0.556</td>
<td>0.148</td>
<td>0.556</td>
<td>Medium</td>
</tr>
<tr>
<td>P22R X</td>
<td>Red direct-view, projection</td>
<td>Y(_2)O(_2):Eu</td>
<td>65</td>
<td>0.640</td>
<td>0.340</td>
<td>0.441</td>
<td>0.528</td>
<td>Medium</td>
</tr>
<tr>
<td>P-22G X</td>
<td>Green direct-view, projection</td>
<td>Zn(_{1-x})Cd(_x)S:Cu,Al</td>
<td>180</td>
<td>0.625</td>
<td>0.340</td>
<td>0.429</td>
<td>0.525</td>
<td>Short</td>
</tr>
<tr>
<td>P-22B X</td>
<td>Blue direct-view, projection</td>
<td>ZnS:Ag</td>
<td>25</td>
<td>0.150</td>
<td>0.065</td>
<td>0.172</td>
<td>0.168</td>
<td>Medium</td>
</tr>
</tbody>
</table>

---

**TABLE 5.6** Typical Characteristics of Common Phosphors

<table>
<thead>
<tr>
<th>Color&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Type</th>
<th>Fluorescent</th>
<th>Phosphorescent</th>
<th>Persistence&lt;sup&gt;b&lt;/sup&gt;</th>
<th>Intended Use</th>
</tr>
</thead>
<tbody>
<tr>
<td>P&lt;sub&gt;1&lt;/sub&gt;</td>
<td>YG</td>
<td>YG</td>
<td>M</td>
<td></td>
<td>Oscillography, radar</td>
</tr>
<tr>
<td>P&lt;sub&gt;2&lt;/sub&gt;</td>
<td>YG</td>
<td>YG</td>
<td>M</td>
<td></td>
<td>Oscillography</td>
</tr>
<tr>
<td>P&lt;sub&gt;3&lt;/sub&gt;</td>
<td>YO</td>
<td>YO</td>
<td>M</td>
<td></td>
<td></td>
</tr>
<tr>
<td>P&lt;sub&gt;4&lt;/sub&gt;</td>
<td>W</td>
<td>W</td>
<td>MS</td>
<td></td>
<td>Direct view television</td>
</tr>
<tr>
<td>P&lt;sub&gt;5&lt;/sub&gt;</td>
<td>B</td>
<td>B</td>
<td>MS</td>
<td></td>
<td>Photographic</td>
</tr>
<tr>
<td>P&lt;sub&gt;6&lt;/sub&gt;</td>
<td>W</td>
<td>W</td>
<td>S</td>
<td></td>
<td></td>
</tr>
<tr>
<td>P&lt;sub&gt;7&lt;/sub&gt;</td>
<td>B</td>
<td>Y</td>
<td>MS(B), L(Y)</td>
<td>Radar</td>
<td></td>
</tr>
<tr>
<td>P&lt;sub&gt;8&lt;/sub&gt;</td>
<td>Replaced by P&lt;sub&gt;7&lt;/sub&gt;</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P&lt;sub&gt;9&lt;/sub&gt;</td>
<td>Not registered</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P&lt;sub&gt;10&lt;/sub&gt;</td>
<td>Dark trace screen</td>
<td>VL</td>
<td>Radar</td>
<td></td>
<td></td>
</tr>
<tr>
<td>P&lt;sub&gt;11&lt;/sub&gt;</td>
<td>B</td>
<td>B</td>
<td>MS</td>
<td></td>
<td>Photographic</td>
</tr>
<tr>
<td>P&lt;sub&gt;12&lt;/sub&gt;</td>
<td>O</td>
<td>O</td>
<td>L</td>
<td>Radar</td>
<td></td>
</tr>
<tr>
<td>P&lt;sub&gt;13&lt;/sub&gt;</td>
<td>RO</td>
<td>RO</td>
<td>M</td>
<td>Radar</td>
<td></td>
</tr>
<tr>
<td>P&lt;sub&gt;14&lt;/sub&gt;</td>
<td>B</td>
<td>YO</td>
<td>MS(B), M(YO)</td>
<td>Radar</td>
<td></td>
</tr>
<tr>
<td>P&lt;sub&gt;15&lt;/sub&gt;</td>
<td>UV</td>
<td>G</td>
<td>UV(VS), G(S)</td>
<td>Flying-spot scanner, photographic</td>
<td></td>
</tr>
<tr>
<td>P&lt;sub&gt;16&lt;/sub&gt;</td>
<td>UV</td>
<td>UV</td>
<td>VS</td>
<td>Flying-spot scanner, photographic</td>
<td></td>
</tr>
<tr>
<td>P&lt;sub&gt;17&lt;/sub&gt;</td>
<td>B</td>
<td>Y</td>
<td>S(B), L(Y)</td>
<td>Oscillography, radar</td>
<td></td>
</tr>
<tr>
<td>P&lt;sub&gt;18&lt;/sub&gt;</td>
<td>W</td>
<td>W</td>
<td>M-MS</td>
<td>Projection television</td>
<td></td>
</tr>
<tr>
<td>P&lt;sub&gt;19&lt;/sub&gt;</td>
<td>O</td>
<td>O</td>
<td>L</td>
<td>Radar</td>
<td></td>
</tr>
<tr>
<td>P&lt;sub&gt;20&lt;/sub&gt;</td>
<td>YG</td>
<td>YG</td>
<td>M-MS</td>
<td>Storage tube</td>
<td></td>
</tr>
<tr>
<td>P&lt;sub&gt;21&lt;/sub&gt;</td>
<td>RO</td>
<td>RO</td>
<td>M</td>
<td>Radar</td>
<td></td>
</tr>
<tr>
<td>P&lt;sub&gt;22&lt;/sub&gt;</td>
<td>W(R,G,B)</td>
<td>W(R,G,B)</td>
<td>MS</td>
<td>Tricolor video (television)</td>
<td></td>
</tr>
<tr>
<td>P&lt;sub&gt;23&lt;/sub&gt;</td>
<td>W</td>
<td>W</td>
<td>MS</td>
<td>Direct-view television</td>
<td></td>
</tr>
<tr>
<td>P&lt;sub&gt;24&lt;/sub&gt;</td>
<td>G</td>
<td>G</td>
<td>S</td>
<td>Flying-spot scanner</td>
<td></td>
</tr>
<tr>
<td>P&lt;sub&gt;25&lt;/sub&gt;</td>
<td>O</td>
<td>O</td>
<td>M</td>
<td>Radar</td>
<td></td>
</tr>
<tr>
<td>P&lt;sub&gt;26&lt;/sub&gt;</td>
<td>O</td>
<td>O</td>
<td>VL</td>
<td>Radar</td>
<td></td>
</tr>
<tr>
<td>P&lt;sub&gt;27&lt;/sub&gt;</td>
<td>RO</td>
<td>RO</td>
<td>M</td>
<td>Color TV monitor</td>
<td></td>
</tr>
<tr>
<td>P&lt;sub&gt;28&lt;/sub&gt;</td>
<td>YG</td>
<td>YG</td>
<td>L</td>
<td>Radar</td>
<td></td>
</tr>
<tr>
<td>P&lt;sub&gt;29&lt;/sub&gt;</td>
<td>P2 and P25 stripes</td>
<td></td>
<td>Radar, indicators</td>
<td></td>
<td></td>
</tr>
<tr>
<td>P&lt;sub&gt;30&lt;/sub&gt;</td>
<td>Canceled</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P&lt;sub&gt;31&lt;/sub&gt;</td>
<td>G</td>
<td>G</td>
<td>MS</td>
<td>Oscillography, bright video</td>
<td></td>
</tr>
<tr>
<td>P&lt;sub&gt;32&lt;/sub&gt;</td>
<td>PB</td>
<td>YG</td>
<td>L</td>
<td>Radar</td>
<td></td>
</tr>
<tr>
<td>P&lt;sub&gt;33&lt;/sub&gt;</td>
<td>O</td>
<td>O</td>
<td>VL</td>
<td>Radar</td>
<td></td>
</tr>
<tr>
<td>P&lt;sub&gt;34&lt;/sub&gt;</td>
<td>BG</td>
<td>YG</td>
<td>VL</td>
<td>Radar, oscillography</td>
<td></td>
</tr>
<tr>
<td>P&lt;sub&gt;35&lt;/sub&gt;</td>
<td>G</td>
<td>B</td>
<td>MS</td>
<td>Oscillography</td>
<td></td>
</tr>
<tr>
<td>P&lt;sub&gt;36&lt;/sub&gt;</td>
<td>YG</td>
<td>YG</td>
<td>VS</td>
<td>Flying-spot scanner</td>
<td></td>
</tr>
<tr>
<td>P&lt;sub&gt;37&lt;/sub&gt;</td>
<td>B</td>
<td>B</td>
<td>VS</td>
<td>Flying-spot scanner, photographic</td>
<td></td>
</tr>
<tr>
<td>P&lt;sub&gt;38&lt;/sub&gt;</td>
<td>O</td>
<td>O</td>
<td>VL</td>
<td>Radar</td>
<td></td>
</tr>
<tr>
<td>P&lt;sub&gt;39&lt;/sub&gt;</td>
<td>YG</td>
<td>YG</td>
<td>L</td>
<td>Radar</td>
<td></td>
</tr>
<tr>
<td>P&lt;sub&gt;40&lt;/sub&gt;</td>
<td>B</td>
<td>YG</td>
<td>MS(B), L(YG)</td>
<td>Low repetition rate (P12 and P16)</td>
<td></td>
</tr>
<tr>
<td>P&lt;sub&gt;41&lt;/sub&gt;</td>
<td>UV</td>
<td>O</td>
<td>VS(UV), L(O)</td>
<td>Radar with light trigger</td>
<td></td>
</tr>
</tbody>
</table>

<sup>a</sup>Color: B = blue, P = purple, G = green, O = orange, Y = yellow, R = red, W = white, UV = ultraviolet.

<sup>b</sup>Persistence to 10% level: VS = less than 1 ms, S = 1–10 ms, MS = 10–1 ms, M = 1–100 ms, L = 100 ms–1 s.

tubes, it is more meaningful to measure luminescence in footlambert using a system the response of which matches the eye. In addition to the use of a suitable detector, a number of other parameters must be controlled if meaningful measurements are to be obtained. Most important of these is the total energy to the screen, which includes

- Anode voltage
- Cathode current
- Raster size

Copyright 2005 by Taylor & Francis Group
FIGURE 5.94  Linearity characteristics of five common phosphors. Conditions: $E_{a2} = 25$ kV ($1 \text{ fl} = 3426 \text{ cd/m}^2$).

Under normal CRT operating conditions, the luminescence of phosphors is essentially proportional to the beam current applied. However, when high beam currents are employed, some phosphors saturate and depart from linear behavior. The same effect can be observed in a direct-view color video display in areas of highlight brightness and with electron guns having small spot size. If linearity of the three primary phosphors is not closely matched, noticeable shift in white field color can result in highlight areas. Figure 5.94 shows some typical light output/beam current curves.

**Thermal quenching** is another mechanism that may result in a loss of phosphor efficiency. In most phosphors the energy transitions become less efficient as screen temperature increases. This phenomenon can be quite pronounced in projection systems where the high-power loading can be responsible for a large increase in screen temperature. Figure 5.95 shows some examples of thermal quenching. Quenching is a transient condition, and the screen returns to normal efficiency after being cooled.

**Electron Gun**

The electron gun is basic to the structure and operation of any cathode-ray device, specifically display devices. In its simplest schematic form, an electron gun may be represented by the diagram in Fig. 5.96, which shows a triode gun in cross section. Electrons are emitted by the cathode, which is heated by the filament to a temperature sufficiently high to release the electrons. Because this stream of electrons emits from the cathode as a cloud rather than a beam, and it is necessary to accelerate, focus, deflect, and

FIGURE 5.95  The loss in phosphor efficiency as a display screen heats at high-current operation (rare-earth green).
Electronics Handbook

otherwise control the electron emission so that it becomes a beam and can be made to strike a phosphor at the proper location and with the desired beam cross section.

**Electron Motion**

The properties of motion for an electron in a uniform electrostatic field are obtained from Newton's second law. The velocity of an emitted electron is given by the following:

\[ V = \left( \frac{2eV}{m} \right)^{1/2} \]

(5.9)

where

- \( e = 1.6 \times 10^{-19} \) C
- \( m = 9.1 \times 10^{-28} \) g
- \( V = -Ex \), the potential through which the electron has fallen

When practical units are substituted for the values in the previous equation, the following results:

\[ V = 5.93 \times 10^5 V^{1/2} \text{ m/s} \]

(5.10)

This expression represents the velocity of the electron. If the electron velocity is at the angle \( \theta \) to the potential gradient in a uniform field, the motion of the electron is specified by

\[ y = \frac{Ex^2}{4V_0 \sin^2 \theta} + \frac{x}{\tan \theta} \]

(5.11)

where \( \theta \) = the electron potential at initial velocity

This equation defines a parabola. The electron trajectory is illustrated in Fig. 5.97, in which the following conditions apply:

- \( y_m \) = maximum height
- \( x_m \) = x displacement at the maximum height
- \( a \) = the slope of the curve

**Tetrode Gun**

The tetrode electron gun includes a fourth electrode, illustrated in Fig. 5.98. The main advantage of the additional electrode is improved convergence of the emitted beam.

Nearly all CRT electron guns currently available have indirectly heated cathodes in the form of a small capped nickel sleeve or cylinder with an insulated coiled tungsten heater inserted from the back end. Most heaters operate at 6.3 V AC at a current of 300–600 mA. Low-power heaters are also available that operated at 1.5 V (typically, 140 mA).

The cathode assembly is mounted on the axis of the modulating or control grid cylinder (or simply grid), which is a metal cup of low permeability or stainless steel about 0.5-in in diameter and 0.375–0.5 in long.
The electron trajectory from an electron gun using the parameters specified in the text.

A small aperture on the order of 10 mils diameter is punched or drilled in the cap. The grid \((G_1)\) voltage is usually negative with respect to the cathode \((K)\).

To obtain electron current from the cathode through the grid aperture, there must be another electrode beyond the aperture at a positive potential great enough for its electrostatic field to penetrate the aperture to the cathode surface. Figure 5.99 illustrates a typical **accelerating electrode** \((G_2)\) in relation to the cathode structure. The accelerating electrode may be implemented in any given device in a number of ways. In a simple accelerating lens in which successive electrodes have progressively higher voltages, the electrode may also be used for focusing the electron beam upon the phosphor, in which case it may be designated the **focusing** (or **first**) anode \((A_1)\). This element is usually a cylinder, longer than its diameter and probably containing one or more disk apertures.

**Electron Beam Focusing**

The general principles involved in focusing the electron beam can be best understood by initially examining optical lenses and then establishing the parallelism between them and electrical focusing techniques.

Figure 5.100 shows a simplified diagram of an electrostatic lens. An electron emitted at zero velocity enters the \(V_1\) region and moves in that region at a constant velocity (because the region has a constant potential). The velocity of the electron in that region is defined by Eq. (5.10) for the straightline component,
with $V_1$ replacing $V$ in the equation. After passing through the surface into the $V_2$ region, the velocity changes to a new value determined by $V_2$. The only component of the velocity that is changed is the one normal to the surface, so that the following conditions are true:

$$V_1 = V_1 \sin I_1$$
$$V_2 \sin I_1 = V_2 \sin I_2$$

*Snell’s law*, also known as the law of refraction, has the form:

$$N_1 \sin I_1 = N_2 \sin I_2$$

where

- $N_1$ = index of refraction for the first medium
- $N_2$ = index of refraction for the second medium
- $I_1$ = angle of the incident ray with the surface normal
- $I_2$ = angle of the refracted ray with the surface normal

The parallelism between the optical and the electrostatic lens is apparent if the appropriate substitutions
are made,

\[ V_1 \sin I_1 = V_2 \sin I_2 \]
\[ \frac{\sin I_1}{\sin I_2} = \frac{V_2}{V_1} \]

For Snell’s law, the following applies:

\[ \frac{\sin I_1}{\sin I_2} = \frac{N_2}{N_1} \]

The analogy between the optical lens and the electrostatic lens is, thus, apparent. The magnification of the electrostatic lens is given by the following:

\[ m = \left( \frac{V_1}{V_2} \right)^2 \frac{S_2}{S_1} \]

The condition of a thin, unipotential lens, where \( V_1 = V_2 \), is illustrated in Fig. 5.101. The following applies:

\[ m = \frac{h_2}{h_1} = \frac{f_2}{X_1} = \frac{-X_2}{f_2} \]

The shape of the electron beam under the foregoing conditions is shown in Fig. 5.102. If the potential at the screen is the same as the potential at the anode, the crossover is imaged at the screen with the magnification.
given by

\[ m = \frac{X_2}{X_1} \]

The magnification can be controlled by changing this ratio, which in turn changes the size of the spot. This is one way to control the quality of the focus. Although the actual lens may not be thin, and in general is more complicated than is shown in the figure, the illustration is sufficient to understand the operation of electrostatic focus.

The size of the spot can be controlled by changing the ratio of \( V_1 \) to \( V_2 \) or of \( X_2 \) to \( X_1 \) in the previous equations. Because \( X_1 \) and \( X_2 \) are established by the design of the CRT, the voltage ratio is the parameter available to the circuit designer to control the size or focus of the spot. It is by this means that focusing is achieved for CRTs using electrostatic control.

### 5.6.3 Color CRT Devices

The shadow-mask CRT is the most common type of color display device. As illustrated in Fig. 5.103, it utilizes a cluster of three electron guns in a wide neck, one gun for each of the colors red, green, and blue. All of the guns are aimed at the same point at the center of the shadow-mask, which is an iron-alloy grid with an array of perforations in triangular arrangement, generally spaced 0.025-in between centers for entertainment television. For high-resolution studio monitor or computer graphic monitor applications, color CRTs with shadow mask aperture spacing of 0.012-in center-to-center or less are readily available.

This triangular arrangement of electron guns and shadow-mask apertures is known as the delta-gun configuration. Phosphor dots on the faceplate just beyond the shadow mask are arranged so that after passing through the perforations, the electron beam from each gun can strike only the dots emitting one color.

All three beams are deflected simultaneously by a single large-diameter deflection yoke, which is usually permanently bonded to the CRT envelope by the tube manufacturer. The three phosphors together are designated P-22, individual phosphors of each color being denoted by the numbers P-22R, P-22G, and P-22B. Most color CRTs are constructed with rare-earth-element-activated phosphors, which offer superior color and brightness compared with previously used phosphors.

Because of the close proximity of the phosphor dots to each other and the strict dependence on angle of penetration of the electrons through the apertures, tight control over electron optics must be maintained. Close attention is also paid to shielding the CRT from extraneous ambient magnetic fields and to degaussing of the shield and shadow mask (usually carried out automatically when the equipment is switched on or off).

Even if perfect alignment of the mask and phosphor triads is assumed, the shadow-mask CRT is still subject to certain limitations, mainly in regard to resolution and luminance. The resolution restriction is the result of the necessity for aligning the mask apertures and the phosphor dot triads; the mask aperture size controls the resolution that can be attained by the device.

Electron beam efficiency in a shadow-mask tube is low, relative to a monochrome CRT. Typical beam efficiency is 9%; considering the three beams of the color tube, total efficiency is approximately 27%. By comparison, a monochrome tube may easily achieve 80% electron beam efficiency. This restriction leads to a significant reduction in luminance for a given input power for the shadow-mask CRT.

### Parallel-Stripe Color CRT

The parallel-stripe class of CRT, such as the popular Trinitron, incorporates fine stripes of red-, green-, and blue-emitting phosphors deposited in continuous lines repetitively across the faceplate, generally in a vertical orientation. (Trinitron is a registered trademark of Sony.) The Trinitron is illustrated in Fig. 5.104. This device, unlike a shadow-mask CRT, uses a single electron gun that emits three electron beams across a diameter perpendicular to the orientation of the phosphor stripes. This type of gun is called the in-line gun. Each beam is directed to the proper color stripe by means of the internal beam-aiming structure and a slitted aperture grille.
FIGURE 5.103 Basic concept of a shadow-mask color CRT: (a) overall mechanical configuration, (b) delta gun arrangement on the tube base, (c) shadow-mask geometry.

The Trinitron phosphor screen is built in parallel stripes of alternating red, green, and blue elements. A grid structure placed in front of the phosphors, relative to the CRT gun, is used to focus and deflect the beams to the appropriate color stripes. Because the grid spacing and stripe width can be made smaller than the shadow-mask apertures and phosphor dot triplets, higher resolutions may be attained with the Trinitron system.
FIGURE 5.104 Basic concept of the Trinitron color CRT: (a) overall mechanical configuration, (b) in-line gun arrangement on the tube base, (c) mask geometry.

Elimination of conventional mask transmission loss, which reduces the electron beam-to-luminance efficiency of the shadow mask tube, permits the Trinitron to operate with significantly greater luminance output for a given beam input power.

The in-line gun is directed through a single lens of large diameter. The tube geometry minimizes beam focus and deflection aberrations, greatly simplifying convergence of the red, green, and blue beams on the phosphor screen.

Basics of Color CRT Design

The shadow-mask CRT is the workhorse of the video display industry. Used in the majority of color video displays since the introduction of color television in the early 1950s, the shadow-mask technique has been refined to yield greater performance and lower manufacturing cost.

Figure 5.105 illustrates the shadow-mask geometry for a tube at face center using in-line guns and a shadow mask of round holes. As an alternative, the shadow-mask may consist of vertical slots, as shown in Fig. 5.106. The three guns and their undeflected beams lie in the horizontal plane. The beams are
shown converged at the mask surface. The beams may overlap more than one hole and the holes are encountered only as they happen to fall in the scan line. By convention, a beam in the figure is represented by a single straight line projected backward at the incident angle from an aperture to an apparent center of deflection located in the deflection plane. In Fig. 5.105, the points B', G', and R', lying in the deflection plane, represent such apparent centers of deflections for blue, green, and red beams striking an aperture under study. (These deflection centers move with varying deflection angles.) Extending the rays forward to the facepanel denotes the printing location for the respective colored dots (or stripes) of a tricolor group. Thus, centers of deflection become color centers with a spacing \( S \) in the deflection plane. The distance \( S \) projects in the ratio \( Q/P \) as the dot spacing within the trio. Figure 5.105 also shows how the mask hole horizontal pitch \( b \) projects as screen horizontal pitch in the ratio \( L/P \). The same ratio applies for projection of mask vertical pitch \( a \). The \( Q \)-space (mask to panel spacing) is optimized to obtain the largest possible theoretical dots without overlap. At panel center, the ideal screen geometry is then a mosaic of equally spaced dots (or stripes).

The stripe screen shown in Fig. 5.106 is used extensively in color CRTs. One variation of this stripe (or line) screen uses a cylindrical faceplate with a vertically tensioned grill shadow-mask without tie bars. Prior to the stripe screen, the standard construction was a tridot screen with a delta gun cluster as shown in Fig. 5.107.

**Guard Band**

The use of guard bands is a common feature for aiding purity in a CRT. The guard band, where the lighted area is smaller than the theoretical tangency condition, may be either positive or negative. In Fig. 5.105, the leftmost red phosphor exemplifies a positive guard band; the lighted area is smaller than the
actual phosphor segment, accomplished by mask hole diameter design. Figure 5.106, on the other hand, shows negative guard band (NGB) (or window-limited) construction for stripe screens. Vertical black stripes about 0.1 mm (0.004-in) wide separate the phosphor stripes forming windows to be lighted by a beam wider than the window opening by about 0.1 mm. Figure 5.107 shows NGB construction of a tridot screen.

**Electron Gun Classifications**

Electron guns for color tubes can be classified according to the main lens configuration, which include:

- Unipotential
- Bipotential
- Tripotential
- Hybrid lenses

The unipotential gun, discussed previously in this chapter as it applies to monochrome operation, is the simplest of all designs. This type of gun is rarely used for color applications, except for small-screen sizes. The system suffers from a tendency toward arcing at high anode voltage, and relatively large low current spots.

The bipotential lens is the most commonly used gun in shadow-mask color tubes. The arrangement of gun electrodes is shown in Fig. 5.108 along with a computer-generated plot of equipotential lines and electron trajectories. The main lens of the gun is formed in the gap between grid 3 and grid 4. When grid 3 operates at 18–22% of the grid-voltage, the lens is referred to as a *low bipotential* configuration, often called simply LoBi. When grid 3 operates at 26–30% of the grid-4 voltage, the lens is referred to as a *high bipotential* or HiBi configuration.
FIGURE 5.107  Delta-gun, round-hole mask negative guard band tridot screen. The taper on the mask holes is shown in the detail drawing only.

The LoBi configuration has the advantages of a short grid 3 and shorter overall length, with parts assembly generally less critical than the HiBi configuration. However, with its shorter object distance (grid-3 length), the lens suffers from somewhat larger spot size than the HiBi. The HiBi, on the other hand, with a longer grid-3 object distance, has improved spot size and resolution. The focus voltage supply for the LoBi also can be less expensive.

It can be seen in the plot Fig. 5.108 that the bipotential beam starts with a crossover of the beam near the cathode, rising to a maximum diameter in the lens. After a double bending action in the lens region the rays depart in a convergent attitude toward the screen of the tube.

Further improvement in focus characteristics has been achieved with a tripotential lens. The electrode arrangement is shown in the computer model of Fig. 5.109. The lens region has more than one gap and requires two focus supplies, one at 40% and the other at 24% of the anode potential. With this refinement the lens has lower spherical aberration. Together with a longer object distance (grids 3–5), the resulting spot size at the screen is smaller than the bipotential designs. Drawbacks of the tripotential gun include

- The assembly is physically longer.
- It requires two focus supplies.
- It requires a special base to deliver the high focus voltage through the stem of the tube.

Hybrid Lenses

Improved performance may be realized by combining elements of the unipotential and bipotential lenses in series. The more common of these configurations are known as UniBi and BiUni. The UniBi structure (sometimes referred to as quadripotential focus) combines the HiBi main lens gap with the unipotential type of lens structure to collimate the beam. As shown in Fig. 5.110, the grid-2 voltage is tied to a grid 4 inserted in the object region of the gun, causing the beam bundle to collimate to a smaller diameter in the main lens. With this added focusing, the gun is slightly shorter than a bipotential gun having an equal focus voltage.

The BiUni structure, illustrated in Fig. 5.111, achieves a similar beam collimation by tying the added element to the anode, rather than to grid 2. The gun structure is shorter because of the added focusing early in the device. With three high gradient gaps, arcing can be a problem in the BiUni configuration.
Trinitron
The Trinitron gun consists of three beams focused through the use of a single large main focus lens of unipotential design. Figure 5.112 shows the three in-line beams mechanically tilted to pass through the center of the lens, then reconverged toward a common point on the screen. The gun is somewhat longer than other color guns, and the mechanical structuring of the device requires unusual care and accuracy in assembly.

Gun Arrangements
Prior to 1970 most color gun clusters used a delta arrangement. Since that time the design trend has been toward the three guns in-line on the horizontal axis of the tube.

The electron optical performance of a delta gun is superior to in-line, in its basic form, by virtue of a larger lens diameter, resulting from more efficient use of the available area inside the neck. Larger neck diameters improve lens diameter, but at the expense of higher deflection power and more difficult convergence problems over the face of the tube. Typical delta gun lens diameters include

- 12 mm (0.5 in) for a 51-mm (2-in) neck OD
- 9 mm (0.4 in) for a 36-mm (1.4-in) neck
- 7 mm (0.3 in) for a 29-mm (1.1-in) neck

Delta guns use individual cylinders to form the electron lenses, and the three guns are tilted toward a common point on the screen. The individual cylinders are subject to random errors in position that can

Copyright 2005 by Taylor & Francis Group
cause misconvergence of the three undeflected spots. Further, carefully tailored current waveforms are applied to magnetic pole pieces on the separate guns to dynamically converge the three beams over the full face of the tube.

In-line guns enjoy one major advantage that more than offsets their less efficient use of available neck area. That is, with three beams on the horizontal axis, a deflection yoke can be built that will maintain dynamic convergence over the full face of the tube, without the need for correcting waveforms. With this major simplification in needed circuitry almost all guns now being built for commercial video applications are of the in-line design. In-line guns are available for 36-, 29-, and 22.5-mm neck diameters (among others). Typical lens diameters for these three guns are 7.5, 5.5, and 3.5 mm, respectively.

**Unitized Construction**

Most in-line guns are of unitized construction in which the three apertures or lens diameters are formed from a single piece of metal. With this arrangement the apertures and lens diameters are accurately fixed with respect to each other so that beam landings at the screen are more predictable than with the cylinder guns. Also, self-converging deflection yokes need the more accurate positioning of the beams in the yoke area. The only trio of gun elements not electrically tied together are the cathodes. Therefore, all varying voltages controlling luminance and color must be applied to the cathodes only. It should be noted that the three beams in unitized guns travel parallel until they reach the final lens gap. Here, an offset, or tilted, lens on the two outboard beams bends the two outer beams toward the center beam at the screen. Any change in the strength of this final lens gap, such as a focus voltage adjustment, will cause a slight change in the undeflected convergence pattern at the screen.

**Guns for High-Resolution Applications**

Improved versions of both the delta and in-line guns noted previously are used in high-resolution data display applications. In both cases the guns are adjusted for the lower beam current and higher resolution needed in data and/or graphics display. The advantages and disadvantages noted in earlier sections for delta and in-line guns apply here also. For example, the use of a delta-type cylinder, or barrel-type gun, requires as many as 20 carefully tailored convergence waveforms to obtain near-perfect convergence over the full face of the tube. The in-line gun, with a self-converging yoke, avoids the need for these waveforms, at the expense of slightly larger spots, particularly in the corners where overfocused haze tails can cause problems.

Figure 5.113 compares spot sizes, at up to 1 mA of beam current, for high-resolution designs compared with commercial receiver-type devices. Both delta and in-line 13- and 19-in vertical (13 V and 19 V)
devices are shown. Note the marked improvement in spot size at current levels below 500 \( \mu A \) for the high-resolution devices.

**Deflecting Multiple Electron Beams**

Deflection of the electron beams in a color CRT represents a difficult technical exercise. The main problems that occur when the three beams are deflected by a common deflection system are associated with the spot distortions that occur in single-beam tubes. However, the effect is intensified by the need for the three beams to cross over and combine as a spot on the shadow mask. The two most significant effects are curvature of field and astigmatism.

Misalignment and misregistration of the three beams of the color CRT will lead to loss of purity for colors produced by combinations of the primary colors. Such reproduction distortions can also result in a reduction in luminance output because a smaller part of the beams are passing through the apertures. Additional errors that must be considered include

- Deflection-angle changes in the yoke-deflection center
- Stray electromagnetic fields

**Defining Terms**

**Cathodoluminescence:** The property of luminescent crystals (phosphors) to emit visible light with bombarded by electrons.

**Delta gun:** In a color CRT, a triangular arrangement of electron guns and shadow-mask apertures positioned so as to permit individual excitation of red, green, and blue phosphor elements.

**Fluorescence:** The property of light emission from a crystal during excitation by electrons.
Guard band: A design technique for a color CRT intended to improve the purity performance of the device by making the lighted area of the screen smaller than the theoretical tangency condition of the device geometry.

In-line gun: In a color CRT, a linear arrangement of electron guns and shadow-mask apertures positioned so as to permit individual excitation of red, green, and blue phosphor elements.

Phosphorescence: The property of light emission from a crystal for some period of time immediately after excitation by electrons is removed.

Thermal quenching: The undesirable mechanism of a phosphor in a CRT device in which phosphor efficiency decreases as the temperature of the phosphor screen increases.

Trinitron: Trade name for a class of color CRTs using in-line guns and in-line shadow-mask/phosphor geometry. Trinitron is a registered trademark of Sony.

References


EIA. 1975. CRTs: Glossary of Terms and Definitions. Publication TEP92, Electronic Industries Association, Washington, DC.


Further Information

The SPIE (the international optical technology society) offers a wide variety of publications on the topic of display systems engineering. The organization also holds technical seminars on this and other topics several times each year. SPIE is headquartered in Bellingham, WA. The Society for Information Display (SID) also holds regular technical seminars and offers a number of publications on the topic of display systems. SID is headquartered in Santa Ana, CA. Furthermore, two books on the topic of display systems are recommended:

The Electronic Industries Association (EIA), Washington, DC, also offers several publications on the topic of display systems as they relate to consumer products.

5.7 Projection Systems

Jerry C. Whitaker

5.7.1 Introduction

The use of large-screen projection displays continues to grow at a rapid rate as the need to present high-resolution video and graphics information steadily increases. High definition television (HDTV) will requires large screens (greater than 40-in diagonal) to provide effective presentation. Some form of projection is, therefore, the only practical solution. The role of HDTV and film in future theaters is also being explored, along with performance criteria for effective large-screen video presentations.

New projection system hardware is taking advantage of liquid crystal (LC) and thin film transistor (TFT) technology originally developed for direct-view flat panel displays. Also, new deformable-membrane light valves used in Schlieren systems are being developed to update oil film light-valve projectors, which have been the mainstay of large-screen projection technology for many years.

The number of specific applications for large-screen displays continues to grow. Military command centers provide a stringent test where high resolution, brightness, and reliability are critical. Flight simulation presents perhaps some of the greatest challenges, as large images must be presented on domes with high brightness, contrast, and resolution. Success in this application may lead to new entertainment vehicles, where domed entertainment theaters will combine traditional presentations with dynamic, high-definition video enhancements.

Extensive developmental efforts are being conducted into large-screen display systems. Much of this research is aimed at advancing new technologies such as plasma, electroluminescent and LCD, lasers, and new varieties of CRTs. Because this section is concerned with vacuum tube technologies, this chapter will focus on CRT-based projection systems. Light valve and laser technologies will also be examined. Solid-state projection systems are covered in Chapter 41.

Display Types

Large-screen projectors fall into four broad classes, or grades:

- **Graphics.** Graphics projectors are the highest quality, and generally most expensive, projectors. These systems are capable of the highest operating frequency and resolution. They can genlock to almost any computer or image source and typically offer resolutions of $2000 \times 2000$ pixels and greater with horizontal sweep rates up to 89 kHz or more.

- **Data.** Data projectors are less expensive than graphics projectors and are suitable for use with common computer image generators, such as PCs equipped with EGA, VGA, and Macintosh standard graphics cards.

- **HDTV.** HDTV projectors provide the quality level necessary to take full advantage of high-definition imaging systems. Resolution on the order of 1125 TV lines is provided at an aspect ratio of 16:9.

- **Video.** Video projectors provide resolution suitable for National Television Systems Committee (NTSC) level images. Display performance of 380–480 TV lines is typical.

Computer signal sources can follow many different, and sometimes incompatible, standards. Multisync projectors are common. Generally speaking, projectors are downward compatible. In other words, most graphic projectors can function as data projectors, HDTV projectors, and video projectors; most data projectors can also display HDTV and video, and so on.

Displays for HDTV Applications

The most significant difference between a conventional display and one designed for high-definition video is the increased resolution and wider aspect ratio of HDTV. Three technologies are practical for viewing high-definition images of 40-in diagonal and larger:

- **Light valve projection display.** Capable of modulating high-power external light sources, light valves are mainly used for large screen displays.

- **CRT projection displays.** Widely use for middle-sized screen displays, CRT projection systems are popular because of their relative ease of manufacturing (and hence competitive cost) and good performance. These displays are likely to be the mainstay technology for HDTV in the near-term future.

- **Flat panel plasma display panel (PDP) and liquid crystal display (LCD).** Flat panel PDP displays are available demonstrated in 40-in and larger sizes.
### TABLE 5.7 Performance Requirements for Large-Screen Military Display Applications

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Performance Requirement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Output luminancea:</td>
<td></td>
</tr>
<tr>
<td>Small conference room</td>
<td>50–200 lm</td>
</tr>
<tr>
<td>Large conference room</td>
<td>200–600 lm</td>
</tr>
<tr>
<td>Control center</td>
<td>600–2000 lm</td>
</tr>
<tr>
<td>Addressable resolution</td>
<td>1280 × 1024 pixels</td>
</tr>
<tr>
<td>Visual resolution</td>
<td>15% absolute minimum modulation depth to an alternate pixel input</td>
</tr>
<tr>
<td>Video interface</td>
<td>15.75–80 kHz horizontal scan, 60–120 Hz noninterlaced refresh</td>
</tr>
<tr>
<td>Video sources</td>
<td>10–20 different sources, autosynchronous lock</td>
</tr>
<tr>
<td>Response to new update</td>
<td>Less than 1 s, no smear to dynamic response</td>
</tr>
<tr>
<td>Reliability</td>
<td>1000 h MTBF, excluding consumables</td>
</tr>
<tr>
<td>Operating cost</td>
<td>Less than $10/h per projector, including AC power</td>
</tr>
<tr>
<td>Maintenance</td>
<td>Level 5 maintenance technician</td>
</tr>
</tbody>
</table>

*aAmbient lighting of 20–40 fc is assumed. The values shown are modulated luminance, spatially averaged over the full screen area. (Source: Blaha, R.J. 1990. Large screen display technology assessment for military applications. Large Screen Projection Displays II, ed. W.P. Bleha, Jr., pp. 80–92. Proc. SPIE 1255.)*

### Displays for Military Applications

Full-color, large-screen display systems can enhance military applications that require group presentation, coordinated decisions, or interaction between decision makers. Projection display technology already plays an important role in operations centers, simulation facilities, conference rooms, and training centers. Each application requires unique values of luminance, resolution, response time, reliability, and video interface. The majority of large-screen applications involve fixed environments where commanders and their staff interact as a group to track an operation and make mission-critical decisions. As sensor technology improves and military culture drives toward scene presence, future applications will require large-screen display systems that offer greater realism and real-time response. Table 5.7 lists the basic requirements for military display applications. The available technologies meeting these demands are illustrated in Fig. 5.114.

#### 5.7.2 Projection System Fundamentals

Video projection systems provide a method of displaying a much larger image than can be generated on a direct-view cathode-ray picture tube. Optical magnification and other techniques are employed to throw an expanded image on a passive viewing surface that may have a diagonal dimension of 75-in or more. Direct-view CRTs are generally restricted to less than 39-in diagonal. A primary factor limiting CRT size is glass strength. To withstand atmospheric pressure on the evacuated envelope, CRT weight increases exponentially with linear dimension.

The basic elements of a projection system, illustrated in Fig. 5.115, include:

- Viewing screen
- Optical elements
- Image source
- Drive electronics

The major differences of projection systems from direct-view displays are embodied in the first three areas, whereas the electronics assembly is (typically) essentially the same as for direct-view systems.

#### Projection Requirements

To provide an acceptable image, a projection system must approach or equal the performance of a direct-view device in terms of brightness, contrast, and resolution. Whereas the first two parameters may be compromised to some extent, large displays must excel in resolution because of the tendency of viewers...
to be positioned less than the normal relative distance of four to eight times the picture height from the viewing surface. Table 5.8 provides performance levels achieved by direct-view video displays and conventional film theater equipment.

Evaluation of overall projection system brightness $B$, as a function of its optical components, can be calculated using the following equation:

$$B = \frac{L_G GTR^M D}{4W_f (f/N)^2(1 + m)^2}$$

where
- $L_G = \text{luminance of the green source (CRT or other device)}$
- $G = \text{screen gain}$
- $T = \text{lens transmission}$

**FIGURE 5.114** Basic grading of full-color display technologies vs. display area for military applications.

**FIGURE 5.115** Principal elements of a video projection system.
TABLE 5.8 Performance Levels of Video and Theater Displays

<table>
<thead>
<tr>
<th>Display System</th>
<th>Luminous Output (brightness), nits, fL</th>
<th>Contrast Ratio at Ambient Illumination, fc</th>
<th>Resolution (TVL)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Television receiver</td>
<td>200–400, 60–120</td>
<td>30:1 at 5</td>
<td>275</td>
</tr>
<tr>
<td>Theater (film projector)</td>
<td>34–69, 10–20(^a)</td>
<td>100:1 at 0.1(^b)</td>
<td>4800</td>
</tr>
</tbody>
</table>

\(^a\)U.S. standard (PH-22.124-1969); see Luxenberg and Kuehn, 1968, p. 29.
\(^b\)Limited by lens flare.

\[ R = \text{mirror reflectance} \]
\[ M = \text{number of mirrors} \]
\[ D = \text{dichroic efficiency} \]
\[ W_G = \text{green contribution to desired white output, } \% \]
\[ f/N = \text{lens } f\text{-number} \]
\[ m = \text{magnification} \]

For systems in which dichroics or mirrors are not employed, those terms drop out.

Two basic categories of viewing screens are employed for projection video displays. As illustrated in Fig. 5.116, the systems are

- **Front projection**, where the image is viewed from the same side of the screen as that on which it is projected.
- **Rear projection**, where the image is viewed from the opposite side of the screen as that on which it is projected.

Front projection depends on reflectivity to provide a bright image, whereas rear projection requires high transmission to achieve that characteristic. In either case, screen size influences display brightness inversely as follows:

\[ B = \frac{L}{A} \]

where
\[ B = \text{apparent brightness, cd/m}^2 \]
\[ L = \text{projector light output, lm} \]
\[ A = \text{screen viewing area, m}^2 \]

Thus, for a given projector luminance output, viewed brightness varies in proportion to the reciprocal of the square of any screen dimension (width, height, or diagonal). An increase in screen width from the...
conventional aspect ratio of 4:3 (1.33) to an HDTV ratio of 16:9 (1.777) requires an increase in projector light output of approximately 33% for the same screen brightness.

To improve apparent brightness, directional characteristics can be designed into viewing screens. This property is termed \( \text{screen gain} G \), and the preceding equation becomes

\[
B = G \times \frac{L}{A}
\]

Gain is expressed as screen brightness relative to a \text{lambertian surface}. Table 5.9 lists some typical front-projection screens and their associated gains.

Screen contrast is a function of the manner in which ambient illumination is treated. Figure 5.116 illustrates that a highly reflective screen (used in front projection) reflects ambient illumination as well as the projected illumination (image). The reflected light thus tends to dilute contrast, although highly directional screens diminish this effect. A rear-projection screen depends on high transmission for brightness but can capitalize on low reflectance to improve contrast. A scheme for achieving this is equivalent to the black matrix utilized in tricolor CRTs. Illustrated in Fig. 5.117, the technique focuses projected light through lenticular lens segments onto strips of the viewing surface, allowing intervening areas to be coated with a black (non-reflective) material. The lenticular segments and black stripes normally are oriented in the vertical dimension to broaden the horizontal viewing angle. The overall result is a screen that transmits most of the light (typically, 60%) incident from the rear while absorbing a large percentage of the light (typically, 90%) incident from the viewing side, thus providing high contrast.

Rear-projection screens usually employ extra elements, including diffusers and directional correctors, to maximize brightness and contrast in the viewing area.

As with direct-view CRT screens, resolution can be affected by screen construction. This is not usually a problem with front-projection screens, although granularity or lenticular patterns can limit image detail. In general, any screen element, such as the matrix arrangement described previously, that quantizes the image

---

**TABLE 5.9 Screen Gain Characteristics for Various Materials**

<table>
<thead>
<tr>
<th>Screen Type</th>
<th>Gain, cd/m²</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lambertian (flat-white paint, magnesium oxide)</td>
<td>0.85–0.90</td>
</tr>
<tr>
<td>White semigloss</td>
<td>1.5</td>
</tr>
<tr>
<td>White pearlescent</td>
<td>1.5–2.5</td>
</tr>
<tr>
<td>Aluminized</td>
<td>1–12</td>
</tr>
<tr>
<td>Lenticular</td>
<td>1.5–2</td>
</tr>
<tr>
<td>Beaded</td>
<td>1.5–3</td>
</tr>
<tr>
<td>Ektalite (Kodak)</td>
<td>10–15</td>
</tr>
<tr>
<td>Scotch-light (3 M)</td>
<td>Up to 200</td>
</tr>
</tbody>
</table>

---

*Copyright 2005 by Taylor & Francis Group*
(breaks it into discrete segments) limits resolution. For 525- and 625-line video, this factor does not provide the limiting aperture. High-resolution applications, however, may require attention to this parameter.

**Projectors for Cinema Applications**

Screen brightness is a critical element in providing an acceptable HDTV large-screen display. Without adequate brightness, the impact on the audience is reduced. Theaters have historically utilized front projection systems for 35-mm film. This arrangement provides for efficient and flexible theater seating. Large-screen HDTV is likely to maintain the same arrangement. Typical motion picture theater projection system specifications are as follows:

- Screen width 30 ft
- Aspect ratio 2.35:1
- Contrast ratio 300:1
- Screen luminance 15 fL
- Center-to-edge brightness uniformity 85%

These specifications meet the expectations of motion picture viewers. It follows that for HDTV projectors to be competitive in theatrical display applications, they must meet similar specifications.

**Optical Projection Systems**

Both refractive and reflective lens configurations have been used for the display of a CRT raster on screens 40-in or more in diagonal width. The first attempts merely placed a lenticular Fresnel lens, or an inefficient $f/1.6$ projection lens in front of a shadow mask direct view tube, as shown in Fig. 5.118(a). The resulting brightness of no greater than 2 or 3 fL was suitable for viewing only in a darkened room. Figure 5.118(b) shows a variation on this basic theme. Three individual CRTs are combined with cross reflecting mirrors and then focused onto the screen. The in-line projection layout is shown in Fig. 5.118(c) using three tubes, each with its own lens. This is the most common system used for multitube displays. Typical packaging to reduce cabinet size for front or rear projection is shown in Fig. 5.118(d) and 5.118(e), respectively.

Because of the off-center positioning of the outboard color channels, the optical paths differ from the center channel, and keystone scanning height modulation is necessary to correct for differences in optical throw from left to right. The problem, illustrated in Fig. 5.119, is more severe for wide-screen formats.

Variables to be evaluated in choosing among the many schemes include the following:

- Source luminance
- Source area
- Image magnification
- Optical-path transmission
- Light collection efficiency (of the lens)
- Cost, weight, and complexity of components and corrective circuitry

The lens package is a critical factor in rendering a projection system cost effective. The package must possess good luminance collection efficiency (small $f$-number), high transmission, good modulation transfer function (MTF), light weight, and low cost. Table 5.10 compares the characteristics of some available lens complements.

The total light incident upon a projection screen is equal to the total light emerging from the projection optical system, neglecting losses in the intervening medium. The distribution of this light generally is not uniform. Its intensity is less at screen edges than at the center in most projection systems as a result of light ray obliquity through the lens ($\cos^4 \theta$ law) and vignetting effects [Luxenberg and Kuehn 1968].

Light output from a lens is determined by collection efficiency and transmittance, as well as source luminance. Typical figures for these characteristics are

- Collection efficiency: 15–25%
- Transmittance: 75–90%
FIGURE 5.118 Optical projection configurations: (a) single tube, single lens rear projection system; (b) crossed-mirror trinescope rear projection; (c) three tube, three lens rear projection systems; (d) folded optics, front projection with three tubes in-line and a dichroic mirror; (e) folded optics, rear projection.

Collection efficiency is partially a function of the light source, and the figure given is typical for a lambertian source (CRT) and lens having a half-field angle of approximately 25°.

Optical Distortions

Optical distortions are important to image geometry and resolution. Geometry is generally corrected electronically, both for pincushion/barrel effects and keystoning, which result from the fact that the three image sources are not coaxially disposed in the common in-line array.

Resolution, however, is affected by lens astigmatism, coma, spherical aberration, and chromatic aberration. The first three factors are dependent on the excellence of the lens, but chromatic aberration can be minimized by using line emitters (monochromatic) or narrow-band emitters for each of the three image sources. Because a specific lens design possesses different magnification for each of the three primary colors (the index of refraction varies with wavelength), throw distance for each must also be adjusted independently to attain perfect registration.

In determining final display luminance, transmission, reflectance, and scattering by additional optical elements such as dichroic filters, optical-path folding mirrors, or corrective lenses must also be accounted
Three tube in-line array. The optical axis of the outboard (red and blue) tubes intersect the axis of the green tube at screen center. Red and blue rasters show trapezoidal (keystone) distortion and result in misconvergence when superimposed on the green raster, thus requiring electrical correction of scanning geometry and focus.

Dichroics exhibit light attenuations of 5–30% and mirrors can reduce light transmission by as much as 5% each. Front-surface mirrors exhibit minimum absorption and scattering but are susceptible to damage during cleaning. Contrast is also affected by the number and nature of optical elements employed in the projection system. Each optical interface generates internal reflections and scattering, which dilute contrast and reduce MTF amplitude response. Optical coatings may be utilized to minimize these effects, but their contribution must be balanced against their cost.

**Image Devices**

CRTs and light valves are the two most common devices for creating images to be optically projected. Each is available in a multitude of variations. Projection CRTs have historically ranged in size from 1 in (2.5 cm) to 13 in (33 cm) diagonal (diameter for round envelope types). Because screen power must increase in proportion to the square of the magnification ratio, it is clear that faceplate dissipation for CRTs used in projection systems must be extremely high. Electrical-to-luminance conversion efficiency for common video phosphors is on the order of 15% (McKechnie, 1981). A 50-in (1.3-m) diagonal screen at 60 fL requires a 5-in (12.7-cm) CRT to emit 6000 fL, exclusive of system optical losses, resulting in

<table>
<thead>
<tr>
<th>Lens type</th>
<th>Aperture, f/</th>
<th>Image Diagonal, mm</th>
<th>Focal Length, mm</th>
<th>Magnification</th>
<th>Response at 300 TVL, %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Refractive, glass</td>
<td>1.6</td>
<td>196</td>
<td>170</td>
<td>8</td>
<td>33</td>
</tr>
<tr>
<td>Refractive, acrylic</td>
<td>1.0</td>
<td>127</td>
<td>127</td>
<td>10</td>
<td>13</td>
</tr>
<tr>
<td>Schmidt</td>
<td>0.7</td>
<td>76</td>
<td>87</td>
<td>30</td>
<td>15</td>
</tr>
<tr>
<td>Fresnel</td>
<td>1.7</td>
<td>305</td>
<td>300</td>
<td>3</td>
<td>6</td>
</tr>
</tbody>
</table>
a faceplate dissipation of approximately 20 W in a 3-in (7.6-cm) × 4-in (10.2-cm) raster. A practical limitation for ambient air-cooled glass envelopes (to minimize thermal breakage) is 1 mW/mm² or 7.74 W for this size display. Accommodation of this incompatibility must be achieved in the form of improved phosphor efficiency or reduced strain on the envelope via cooling. Because phosphor development is a mature science, maximum benefits are found in the latter course of action with liquid cooling assemblies employed to equalize differential strain on the CRT faceplate. Such implementations produce an added benefit through reduction of phosphor thermal quenching and thereby supply up to 25% more luminance output than is attainable in an uncooled device at equal screen dissipation (Kikuchi et al., 1981).

A liquid-cooled CRT assembly, shown in Fig. 5.120, depends on a large heat sink to carry away and dissipate a substantial portion of the heat generated in the CRT. Large-screen projectors using such assemblies commonly operate CRTs at four to five times their rated thermal capacities. Economic constraints mitigate against the added cost of cooling assemblies, however, and methods to improve phosphor conversion efficiency and optical coupling/transmission efficiencies continue to be investigated.

Concomitant to high power are high voltage and/or high beam current. Each has benefits and penalties. Resolution, dependent on spot diameter, is improved by increased anode voltage and reduced beam current. For a 525- or 625-line display, spot diameter should be 0.006 in (0.16 mm) on the 3-in × 4-in (7.6-cm × 10.2-cm) raster discussed previously. Higher resolving power displays require yet smaller spot diameters, but a practical maximum anode-voltage limit is 30–32 kV when X radiation, arcing, and stray emission are considered. Exceeding 32 kV typically requires special shielding and CRT processing during assembly.

One form of the in-line array benefits from the relatively large aperture and light transmission efficiency of Schmidt reflective optics by combining the electron optics and phosphor screen with the projection optics in a single tube. The principle components of an integral Schmidt system are shown in Fig. 5.121.
Electrons emitted from the electron gun pass through the center opening in the spherical mirror of the reflective optical system to scan a metal-backed phosphor screen. Light from the phosphor (red, green, or blue, depending on the color channel) is reflected from the spherical mirror through an aspheric corrector lens, which serves as the face of the projection tube. Schmidt reflective optical systems are significantly more efficient than refractive systems because of the lower $f$ characteristic and the reduced attenuation by glass in the optical path.

**Discrete-Element Display Systems**

The essential characteristic defining discrete-element display devices is division of the viewing surface (volume for three-dimensional displays) into separate segments that are individually controlled to generate an image. Each element, therefore, embodies a dedicated controlling switch or valve as opposed to raster display devices, which employ one (or a small number of) control device(s) for activation of all display elements.

**Flood Beam CRT**

Large displays (20 m diagonal and larger) have been employed for mass audiences in stadiums to provide special coverage of sporting events. These consist of flood beam CRT arrays in which each device fulfills the function of a single phosphor dot in a delta-gun shadow-mask CRT. Thus, each display element consists of a trio of flood beam tubes, one red, one green, one blue, each with a 1–6 in (2.5–15 cm) diameter. A fully NTSC-capable display would require in excess of 400,000 such tubes (147,000 of each color) to be individually addressed. Practical implementations have employed less than 40,000, thus requiring substantially reduced drive complexity.

**Matrix Addressing**

Provision for individual control of each element of a discrete-element display implies a proliferation of electrical connections and, therefore, rapidly expanding cost and complexity. These factors are minimized in discrete-element displays through matrixing techniques wherein a display element must receive two or more gating signals before it is activated. Geometric arrangements of the matrixing connectors guarantee that only the appropriate elements are activated at a given time.

Matrix addressing requires that the picture elements of a panel be arranged in rows and columns, as illustrated Fig. 5.122 for a simplified $4 \times 4$ matrix. Each element in a row or column is connected to all other elements in that row or column by active devices (typically transistors). Thus, it is possible to select the drive signals so that when only a row or column is driven, no elements in that row or column is activated. However, when both the row and column containing a selected element are driven, the sum of the two signals will activate a given element. The resulting reduction in the number of connections required, compared to a discrete-element drive, is significant because the number of lines is reduced from one-per-element to the sum of the number of rows and columns in the panel.

For active matrix addressing to work properly, it is necessary to have a sharp discontinuity in the transfer characteristic of the device at each pixel at the intersection of the row and column address lines. In some cases, the characteristics of the display device itself can be used to provide this discontinuity. The breakdown potential of a gas discharge can be used (for plasma displays) or alternatively, liquid crystals can be made to have a reasonably sharp transfer curve.

Stadium applications require color video displays that can be viewed in sunlight. In such cases, the high brightness required is achieved with a matrix system. A bright source at each pixel is controlled by its own drive circuit, as illustrated in Fig. 5.123. Light sources are typically either color incandescent light bulbs or electron-excited fluorescent phosphors. The switching matrix must control the intensity of each color in each pixel of the display. For video, the system must have a continuous gray scale. Because of the cost and complexity of providing individual control of each pixel light source, such displays usually have somewhat lower resolution than those using light valves.
5.7.3 Projection Display Systems

The use of large-screen video displays has been increasing for a wide variety of applications. Because of the greater resolution provided by HDTV and advanced computer graphics systems, higher image quality is being demanded from large-screen display systems. Available technologies include the following:

- CRT-based display systems
- *Eidophor* reflective optical system
- *Talaria* transmissive color system
- Dual light valve system for HDTV
- Laser beam projection scanning system
- LCD projection system

HDTV applications present additional, more stringent requirements than conventional 525/625-line color displays, including:

- Increased light output for wide screen presentation
- Increased horizontal and vertical resolution
- Broader gamut of color response to meet system specifications
- Projection image aspect ratio of 16:9 rather than 4:3
CRT Projection Systems

Large-screen color projection systems typically employ three monochrome CRT assemblies (red, green, blue) and the necessary optics to project full-color images. CRT systems based on 5- and 7-in CRT technology are popular because of their low cost; however, they suffer from low luminance levels (usually less than 50–60 average lumens). Higher luminance levels are provided by 9-in CRTs (160 lumens is typical). The larger CRT surface area permits increased beam energy without sacrificing resolution. Still higher luminance and resolution are offered by 13-in CRT projection systems.

Static and dynamic convergence circuitry is used to align the three beams over the display area. Digital convergence circuits permit convergence files tailored for each video source to be stored within a dedicated microprocessor. Advanced operating capabilities include autosync to a variety of input sources, built-in diagnostics, and automatic setup.

With CRT technology, high luminance and high resolution are conflicting requirements. For a given acceleration voltage, increased beam current will provide increased luminance, but reduced resolution because the spot size tends to increase. High-luminance levels also raise the operating temperature of the device, which may shorten its expected lifetime.

Multiple sets of CRT-based projection systems can be linked to increase luminance and resolution for a given application. The multiple beams overlay each other to yield the improved performance. Convergence of the 6–12 CRT assemblies, however, can become quite complex. Multiple CRT systems are satisfactory for NTSC video sources, but are usually inadequate for the display of computer-generated graphics data with single pixel characters.

The mosaic approach to improved luminance and resolution subdivides the screen into segments, each allocated to an individual projection system. As a result, larger images are delivered with higher overall performance. This scheme has been used effectively in simulation training to provide seamless, panoramic...
Light Valve Systems

Light valves may be defined as devices which, like film projectors, employ a fixed light source modulated by an optical-valve intervening the source and projection optics. Light valve display technology is a rapidly developing discipline. Light valve systems offer high brightness, variable image size, and high resolution. Table 5.11 lists some of the more common light valve systems and their primary operating parameters. The first four systems are based on electron-beam addressing in a CRT. The last system is based on liquid crystal light valve (LCLV) technology. Progress in light valve technology for HDTV depends on developments in two key areas:

- Materials and technologies for light control
- Integrated electronic driving circuits for addressing picture elements

**Eidophor Reflective Optical System**

Light valve systems are capable of producing images of substantially higher resolution than are required for 525/625-line systems. They are ideally suited to large screen theater displays.

In a manner similar to film projectors, a fixed light source is modulated by an optical value system (Schlieren optics) located between the light source and the projection optics (see Fig. 5.124). In the basic Eidophor system, collimated light typically from a 2-kW xenon source (component 1 in the figure) is directed by a mirror to a viscous oil surface in a vacuum by a grill of mirrored slits (component 3). The slits are positioned relative to the oil-coated reflective surface so that when the surface is flat, no light is reflected back through the slits. An electron beam scanning the surface of the oil with a video picture raster (components 4–6) deforms the surface in varying amounts, depending on the video modulation of the scanning beam. Where the oil is deformed by the modulated electron scanning beam, light rays from the mirrored slits are reflected at an angle that permits them to pass through the slits to the projection lens. The viscosity of the liquid is high enough to retain the deformation over a period slightly greater than a television field.

Projection of color signals is accomplished through the use of three units, one for each of the red, green, and blue primary colors converged on a screen.

**Talaria Transmissive Color System**

The Talaria system also uses the principle of deformation of an oil film to modulate light rays with video information. However, the oil film is transmissive rather than reflective. In addition, for full-color displays, only one gun is used to produce red, green, and blue colors. This is accomplished in a single light valve by the more complex Schlieren optical system shown in Fig. 5.125.
FIGURE 5.124  Mechanical configuration of the Eidophor projector optical system.

FIGURE 5.125  Functional operation of the General Electric single-gun light valve system.
Colors are created by writing diffraction grating, or grooves, for each pixel on the fluid by modulating the electron beam with video information. These gratings break up the transmitted light into its spectral colors, which appear at the output bars where they are spatially filtered to permit only the desired color to be projected onto the screen.

Green light is passed through the horizontal slots and is controlled by modulating the width of the raster scan lines. This is done by means of a high-frequency carrier, modulated by the green information, applied to the vertical deflection plates. Magenta light, composed of red and blue primaries, is passed through the vertical slots and is modulated by diffraction gratings created at right angles (orthogonal diffraction) to the raster lines by velocity modulating the electron beam in the horizontal direction. This is done by applying 16-MHz and 12-MHz carrier signals for red and blue, respectively, to the horizontal deflection plates and modulating them with the red and blue video signals. The grooves created by the 16-MHz carrier have the proper spacing to diffract the red portion of the spectrum through the output slots while the blue light is blocked. For the 12-MHz carrier, the blue light is diffracted onto the screen while the red light is blocked. The three primary colors are projected simultaneously onto the screen in registry as a full-color picture.

To meet the requirements of HDTV, the basic Talaria system can be modified as shown in Fig. 5.126. In the system, one monochromatic unit with green dichroic filters produces the green spectrum. Because of the high scan rate for HDTV (on the order of 33.75 kHz), the green video is modulated onto a 30-MHz carrier instead of the 12 or 15 MHz used for 525- or 625-line displays. Adequate brightness levels are produced using a 700-W xenon lamp for the green light valve and a 1.3-kW lamp for the magenta (red and blue) light valve.

A second light valve with red and blue dichroic filters produces the red and blue primary colors. The red and blue colors are separated through the use of orthogonal diffraction axes. Red is produced when the writing surface diffracts light vertically. This is accomplished by negative amplitude modulation of a 120-MHz carrier, which is applied to the vertical diffraction plates of the light valve. Blue is produced when the writing surface diffracts light horizontally. This is accomplished by modulating a 30-MHz carrier with the blue video signal and applying it to the horizontal plate, as is done in the green light valve.

The input slots and the output bar system of the conventional light valve are used, but with wider spacing of the bars. Therefore, the resolution limit is increased. The wider bar spacing is achievable because the red and blue colors do not have to be separated on the same diffraction axis as in the single light valve system. This arrangement eliminates the cross-color artifact present with the single light valve system, and therefore improves the overall colorimetric characteristics, as shown in Fig. 5.127.

High-performance electron guns help provide the required resolution and modulation efficiency for HDTV systems of up to 1250 lines. The video carriers have been optimized to increase the signal bandwidth capability to 30 MHz.

In the three-element system, all three devices are monochrome light valves with red, blue, and green dichroic filters. The use of three independent light valves improves color brightness, resolution, and colorimetry. Typically, the three light valves are individually illuminated by xenon arc lamps operating at 1 kW for the green and at 1.3 kW for the red and blue light valves.

The contrast ratio is an important parameter in light valve operation. The amount of light available from the arc lamp is basically constant; the oil film modulates the light in response to picture information. The key parameter is the amount of light that is blocked during picture conditions when totally dark scenes are being displayed (the darkfield performance). Another important factor is the ability of the display device to maintain a linear relationship between the original scene luminance and the reproduced picture. The amount of predistortion introduced by the camera must be compensated for by an opposite amount of distortion at the display device.

**Laser Beam Projection Scanning System**

Several approaches to laser projection displays have been implemented. The most successful employs three optical laser light sources whose coherent beams are modulated electro-optically and deflected by electromechanical means to project a raster display on a screen. The scanning functions are typically
FIGURE 5.126 Functional operation of the two-channel HDTV light valve system.

provided by a rotating polygon mirror and a separate vibrating mirror. A block diagram of the basic system is shown in Fig. 5.128.

The flying spots of light used in this approach are one scan line (or less) in height and a small number of pixels wide. This means that any part of the screen may only be illuminated for a few nanoseconds. The scanned laser light projector is capable of high contrast ratios (as high as 1000:1) in a darkened environment. A laser projector may, however, be subject to a brightness variable referred to as speckle. Speckle is a sparkling effect resulting from interference patterns in coherent light. This effect causes a flat, dull projection surface to look as though it had a beaded texture. This tends to increase the perception of brightness, at the expense of image quality.

Figure 5.129 shows the configuration of a laser projector using continuous-wave lasers and mechanical scanners. The requirements for the light wavelengths of the lasers are critical. The blue wavelength must
be shorter than 477 nm, but as long as possible. The red wavelength should be longer than 595 nm, but as short as possible. Greens having wavelengths of 510, 514, and 532 nm have been used with success. Because laser projectors display intense colors, small errors in color balance can result in significant distortions in gray scale or skin tone. The requirement for several watts of continuous-wave power further limits the usable laser devices. Although several alternatives have been considered, most color laser projectors use argon ion lasers for blue and green, and an argon ion laser pumped dye laser for red.

To produce a video signal, a modulator is required with a minimum operating frequency of 6 MHz. Several approaches are available. The bulk acousto-optic modulator is well suited to this task in high-power laser projection. A directly modulated laser diode is used for some low-power operations. The modulator does not absorb the laser beam, but rather deflects it as needed. The angle of deflection is determined by the frequency of the acoustic wave.
The scanner is the component of the laser projector that moves the point of modulated light across and down the image plane. Several types of scanners may be used including mechanical, acousto-optic, and electro-optic. Two categories of scanning devices are used in the system:

- **Line scanner**, which scans the beam in horizontal lines across the screen. Lines are traced at 15,000–70,000 times per second. The rotating polygon scanner is commonly used, featuring 24–60 mirrored facets. Figure 5.130 illustrates a mechanical rotating polygon scanner.

- **Frame scanner**, which scans the beam vertically and forms frames of horizontal scan lines. The frame scanner cycles 50–120 times per second. A galvonometer-based scanner is typically used in conjunction with a mirrored surface. Because the mirror must fly back to the top of the screen in less than a millisecond, the device must be small and light. Optical element are used to force the horizontally scanned beam into a small spot on the frame scanner mirror.
The operating speed of the deflection components may be controlled either from an internal clock, or locked to an external timebase.

The laser projector is said to have infinite focus. To accomplish this, optics are necessary to keep the scanned laser beam thin. Such optics allow a video image to remain in focus from approximately 4 ft to infinity. For high-resolution applications, the focus is more critical.

There is an undesirable byproduct of high-power laser operation. Most devices are water cooled and use a heat exchanger to dump the waste heat.

Another approach large-screen display employs an electron beam pumped monocrystalline screen in a CRT to produce a 1-in raster. The raster screen image is projected by conventional optics, as shown in Fig. 5.131. This technology promises three optical benefits:

- High luminance in the image plane
- Highly directional luminance output for efficient optical coupling
- Compact, lightweight, and inexpensive projection optics

Full-color operations is accomplished using three such devices, one for each primary color.

![FIGURE 5.131 Laser-screen projection CRT.](image-url)

![FIGURE 5.132 The relationship between misconvergence and display resolution.](image-url)
TABLE 5.12  Appearance of Scan Line Judder in a Projection Display

<table>
<thead>
<tr>
<th>Condition of Adjacent Lines</th>
<th>S/N (p- p)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clearly overlapped</td>
<td>Less than 69 dB</td>
</tr>
<tr>
<td>Just before overlapped</td>
<td>75 dB</td>
</tr>
<tr>
<td>No judder</td>
<td>Greater than 86 dB</td>
</tr>
</tbody>
</table>

5.7.4 Operational Considerations

Accurate convergence is critical to display resolution. Figure 5.132 shows the degradation in resolution resulting from convergence error. It is necessary to keep convergence errors to less than half the distance between the scanning lines to hold resolution loss below 3 dB. Errors in convergence also result in color contours in a displayed image. Estimates have put the detectable threshold of color contours at 0.75–0.5 min of arc. This figure also indicates that convergence error must held under 0.5 scanning lines.

Raster stability influences the short-term stability of the display. The signal-to-noise ratio (S/N) and raster stability relationships for deflection circuits are shown in Table 5.12. A S/N equivalent of 1/5 scanning line is necessary to obtain sufficient raster stability. In HDTV applications, this translates to approximately 80 dB. Other important factors are high speed and improved efficiency of the deflection circuit.

A number of projection manufacturers have begun to incorporate automatic convergence systems into their products. These usually take the form of a charge-coupled device (CCD) camera sensor that scans various portions of the screen as test patterns are displayed.

![Diagram](image)

**FIGURE 5.133** Relationship between screen type and viewing angle: (a) conventional (flat) screen, (b) high gain (curved) screen.
Screen Considerations

Because most screens reflect both ambient light in the room and the incident light from the projector (in a front projection system), lighting levels are an important consideration. A rear screen system can reduce this problem. Ambient light on the screen's projector side reflects harmlessly back toward the projector, away from the audience. Light on the viewing side of the screen may fall on the screen but the human eye has experience in filtering out this kind of reflection. The result is that a rear screen projector provides an increased contrast ratio (all other considerations being equal); it can, therefore, be used in lighter areas. Special screen coatings can also help scatter ambient light away from the viewers, enhancing the rear screen image. Coatings can also increase the apparent brightness of screens operated from the front. The biggest gains, however, are made by using a screen that is curved slightly. Such screens can increase the apparent brightness of an image. Curved screens, however, also narrow the acceptable field of view. High-gain screens provide a typical viewing angle of ±50° from the centerline of the screen. Conventional (low-gain) screens can provide viewing of up to ±90° from the centerline. Figure 5.133 illustrates the tradeoffs involved.

Defining Terms

**Darkfield performance**: The amount of light that is blocked in a light valve projector when a totally dark scene is being displayed. This parameter is critical to good contrast ratio.

**Front projection system**: A projection device where the image is viewed from the same side of the screen as that on which it is projected.

**Infinite focus**: A characteristic of a laser projector in which the video image remains in focus from approximately four feet to infinity.

**Light valve**: A projection device that employs a fixed light source modulated by an optical-valve intervening the source and projection optics.

**Matrix addressing**: The control of a display panel consisting of individual light-producing elements by arranging the control system to address each element in a row/column configuration.

**Rear projection system**: A projection device where the image is viewed from the opposite side of the screen as that on which it is projected.

**Screen gain**: The improvement in apparent brightness relative to a lambertian surface of a projection system screen by designing the screen with certain directional characteristics.

**Speckle**: A sparkling effect in a laser beam projection system resulting from interference patterns in coherent light.

References


EIA. X-radiation measurement procedures for projection tubes. TEPAC Pub. 102, Electronic Industries Association, Washington, DC.


Sears, F.W. 1946. *Principles of Physics, III. Optics*. Addison-Wesley, Cambridge, MA.


Further Information

The SPIE (the international optical technology society) offers a wide variety of publications on the topic of display systems engineering. The organization also holds technical seminars on this and other topics several times each year. SPIE is headquartered in Bellingham, WA. The Society for Information Display (SID) also holds regular technical seminars and offers a number of publications on the topic of projection display systems. SID is headquartered in Santa Ana, CA.
Two books on the topic of display systems are recommended: