3

Incandescent, discharge, and arc lamp sources

DAVID O. WHARMBY

3.1 Overview of sources
3.2 Light production
3.3 Radiation fundamentals
  3.3.1 Full radiator radiation and limits on emission
  3.3.2 Absorption and emittance
  3.3.3 Étendue
  3.3.4 Use of light in systems
  3.3.5 Color properties and color temperature of sources
  3.3.6 Radiation from atoms and molecules in extended sources
3.4 Incandescent lamps
  3.4.1 Emission
  3.4.2 Tungsten–halogen lamps
  3.4.3 Varieties of tungsten–halogen lamps
  3.4.4 Lamps with integral reflectors
  3.4.5 Lamps with IR reflectors
  3.4.6 IR sources
3.5 Discharge lamps with electrodes
  3.5.1 Stable discharge operation of discharges with electrodes
  3.5.2 Electrode regions
3.6 Types of LP discharges
  3.6.1 Low-pressure mercury rare-gas discharges
  3.6.2 Applications of LP mercury discharges
3.7 HP discharges
  3.7.1 General features of HP discharge lamps
  3.7.2 HP metal halide lamps
  3.7.3 Operating principles of metal halide lamps
  3.7.4 Applications of HP discharge lamps
3.8 Electrical characteristics of discharges
  3.8.1 Breakdown and starting in discharge lamps
  3.8.2 Steady state electrical characteristics
  3.8.3 AC operation
3.9 Other methods of excitation of discharges
  3.9.1 Pulsed light sources
  3.9.2 Dielectric barrier discharges
  3.9.3 Excitation by induction and by microwaves
3.10 LEDs from the perspective of conventional lighting
Appendix 3A
References

3.1 Overview of Sources

There is a very wide range of incandescent and discharge lamps. The majority of these are sold as general lighting sources, but many are suited to optoelectronic applications. The major lamp companies, and numerous specialty lamp manufacturers also make lamps for applications other than...
general illumination. Examples of these applications are: projection, video, film, photographic, architectural, entertainment and other special effects, fiber optic illumination including numerous medical and industrial applications, photo-biological processes, photochemical processing, microlithography, solar simulation, sun-tanning, disinfection, ozone generation, office automation, scientific applications, heating etc.

LED sources are covered in detail in Chapter 10. Section 3.10 of this chapter makes some brief comments on the applications in which LEDs are competing with conventional lamps.

This chapter will concentrate on principles and will be illustrated by a number of examples. These principles should make it possible to understand the wealth of information in manufacturer’s websites and catalogues. A selected list of manufacturers is given in Appendix.


3.2 LIGHT PRODUCTION

Most optical radiation is the result of accelerating electrons and causing them to make inelastic collisions with atoms, ions, molecules or the lattice structure of solids. In the UV, visible and near IR, the photons are the result of electronic transitions between energy levels of these materials.

There are exceptions; in synchrotron radiation and related processes emission is from accelerated electrons.

As particle densities increase in the source, the spectral features broaden out until, in incandescent sources the spectrum is continuous. Discharge sources generally emit spectral lines of atoms and molecules that are broadened to an extent depending on the pressure. Lamps of various types therefore emit a wide range of spectral features ranging from narrow atomic lines to a full continuum. The types of spectra are often critical for optical applications [5] (see Appendix—Oriel Instruments for a selection of spectra).

In incandescent lamps, the radiation is from the surface of a hot material. In discharge lamps, conduction is the result of ionization of the gas; any light emission is a volume process. The task of the lamp designer is to ensure that this ionization is also accompanied by copious radiation of the correct quality for the application.

3.3 RADIATION FUNDAMENTALS

3.3.1 Full radiator radiation and limits on emission

Both in incandescent and discharge lamps, electron motion is randomized. In all cases of practical interest, the drift velocity of the electrons in the applied electric field is much less than the mean velocity. An electron energy distribution function is established that can usually be characterized by an electron temperature $T_e$. The distribution function may be far from Maxwellian when particle densities are low, or under transient conditions. It is the electrons in the high-energy tail of the distribution that excite the atoms, with subsequent emission of radiation.

The spectral radiance $L_e(\lambda, T)$ of the full radiator or black body is given by Planck’s equation (Chapter 8, where radiometric and photometric quantities are also defined). The spectral radiance is plotted in Figure 3.1 for temperatures typical of those found in incandescent and discharge lamps. Convenient units for spectral radiance are $\text{W m}^{-2} \text{sr}^{-1} \text{nm}^{-1}$, obtained by multiplying the value of $c_1$ in Chapter 8 by $10^{-9}$.

For incandescent or high-pressure (HP) discharge sources the electron temperature $T_e$ is close in value to the temperature $T$ of the solid or vapor, but for low-pressure (LP) discharges in which collisions between electrons and heavy particles are comparatively rare, $T_e$ may be very much higher than the gas temperature. The Planck equation therefore forms a fundamental limit to the radiance that may be obtained from any source in which the electron motion is randomized. This sets a fundamental limit on the spectral distribution, the energy efficiency and the radiance of the source.
3.3 Radiation fundamentals

3.3.2 Absorption and emittance

For radiation falling on a surface

\[ \alpha(\lambda, T, \theta) + t(\lambda, T, \theta) + r(\lambda, T, \theta) = 1 \]  (3.1)

where the fractions \( \alpha(\lambda, T, \theta) \), \( t(\lambda, T, \theta) \) and \( r(\lambda, T, \theta) \) are known as absorbance, transmittance and reflectance, respectively. In general, they depend on the wavelength, temperature and angle \( \theta \) between a ray and the normal to the surface.

The spectral emittance \( \varepsilon(\lambda, T, \theta) \) is the ratio of the thermal emission from the surface to that of a full radiator (black body) at the same temperature, wavelength and angle. This quantity is also known as spectral emittance. Derived from very general thermodynamic arguments, Kirchhoff’s law [6] states that

\[ \varepsilon(\lambda, T, \theta) = a(\lambda, T, \theta). \]  (3.2)

For a perfect absorber, \( \alpha(\lambda, T, \theta) = 1 \). Therefore, the spectral emittance of a full radiator is unity; a good approximation can be made by forming a cavity from an absorbing material.

All real materials have \( \varepsilon(\lambda, T, \theta) < 1 \). The best characterized material is tungsten (Figure 3.2) [1]. Selective emittance is characteristic of most materials; in metals the emittance tails off at long wavelengths, whereas refractory oxides usually have a region of high emittance in the IR.

3.3.3 Étendue

For all optical systems geometry determines how much of the radiation generated by the source can be used by the optical system. This behavior depends on a very general concept called étendue \( \varepsilon \), also known as geometric extent [6–8].

A definition of étendue is

\[ \varepsilon = \iint \cos \theta \, dA \, d\Omega (m^2 \, sr) \]  (3.3)

where \( \cos \theta \, dA \) is the projected area of the source under consideration, and \( d\Omega \) is the solid angle into which it is radiating. Notice that the units are geometric, with no mention of amounts of radiation. A more general form is used when refractive indices are \( >1 \) [6]. Energy conservation requires that étendue is conserved in a lossless optical system; if there are losses caused by aberrations, scattering, or diffraction, étendue increases through the system. The étendue of a bundle of rays passing through an optical system either stays the same (ideal) or increases, but never decreases.

A simple example demonstrates some of the issues. Imagine projecting an image of the sun onto a surface. The diameter of the sun is about \( 1.4 \times 10^9 \) m with an area \( A_S = 1.5 \times 10^{18} \) m\(^2\). Our distance from the sun is about \( 1.5 \times 10^{11} \) m. Suppose the lens has a focal length of \( f = 100 \) mm and a diameter of 10 mm so that its area \( A_L = 8 \times 10^{-5} \) m\(^2\). The solid angle \( \Omega_0 \) subtended by the lens at the sun is therefore about \( 3.5 \times 10^{-27} \) sr. In this simple geometry the étendue...
\[ \varepsilon = A_S \Omega_0 = 5.4 \times 10^{-9} \text{m}^2 \text{sr} \]

The image is brought to a focus at a distance \( f \) in a converging beam of solid angle \( \Omega_L = A_L / f^2 = 8 \times 10^{-3} \text{sr} \). Assuming a perfect optical system so that étendue is conserved, the image area is therefore \( A_I = \varepsilon / \Omega_L = 7 \times 10^{-7} \text{m}^2 \), giving an image diameter of about 0.5 mm.

If we want to focus the sun onto a smaller spot, a lens of the same area needs to have a shorter focal length. Aberrations in a nonideal lens then cause some of the light to fall outside the area predicted above, increasing étendue. Scattering and diffraction are also losses that increase étendue. In general, the integration in Equation 3.3 has to be done numerically, e.g., by using an optical design code.

Étendue is also the quantity that determines how the power \( \Phi \) (W) in the beam is related to the radiance \( L \) (Wm\(^{-2}\) sr\(^{-1}\)), as inspection of the units will confirm:

\[ \Phi = L \varepsilon t (W) \quad (3.4) \]

\( t \) is the transmittance of the lens (and related optics). Conservation of étendue and of energy means that radiance can never be increased by an optical system.

In a projector, there is always some component that has the smallest (limiting) étendue. Often this will be the film or light gate with its associated projection lens. If the étendue of the source is greater than this, some light will miss the light gate and be wasted. On the other hand, if the étendue of the light gate is much larger than that of the source then the gate will not be fully illuminated. The aim must therefore be to reduce the étendue of the source as far as possible, since it is usually much greater than the limiting étendue. This will minimize the amount of light that misses the light gate. Suppose that a projector lamp has a source of area \( A_S \) that radiates in all directions so that the solid angle is \( 4\pi \) and the source étendue is \( \varepsilon_S = 4\pi A_S \). The limiting étendue \( \varepsilon_L \) of the system will be usually be that of the light gate. In order that \( \varepsilon \) does not greatly exceed \( \varepsilon_L \), with consequent wastage of light, the area of the source must be very small because the source solid angle is so large. Major advances in projector lamps have been to use HP arcs with an arc gap as small as 1 mm (see Section 3.7.4) and an effective area in the region of 0.1 mm\(^2\).

The étendue concept is very general. It applies to any illumination system from fiber optics to street lanterns. For example, one of the benefits of LEDs is that their low étendue allows efficient use of the relatively low radiated fluxes; this is a reason why LED headlights for cars are a possibility.

### 3.3.4 Use of light in systems

The luminous flux in lumens (lm) [3, Chapter 1]

\[ \Phi_\lambda = 683 \int_{380}^{780} \Phi_\lambda V(\lambda) d\lambda \quad (3.5) \]
where \( \Phi_e \lambda \) is the spectral radiant flux in W nm\(^{-1}\) and \( V(\lambda) \) is the spectral luminous efficiency for photopic vision (Chapter 8). The factor 683 (lm W\(^{-1}\)) converts power to luminous flux. It is also useful to define the luminous efficiency of radiation

\[
K = \frac{\Phi_e}{\Phi_e}. \quad (3.6)
\]

The (luminous) efficacy of a source is

\[
\eta_e = \frac{\Phi_e}{P_in} \left( \text{lm W}^{-1} \right). \quad (3.7)
\]

For many commercial lamps, the input power \( P_{in} \) is defined as the power into terminals of the lamp, whereas self-contained sources (such as compact fluorescent lamps), or lamps sold as a system (such as some electrodeless lamps) \( P_{in} \) is taken to be the power coming from the electricity supply \( P_{wall} \). The latter power is greater because it contains the losses in the lamp circuit; users should be aware of this possibility for confusion.

Many lighting systems are driven and controlled by electronics and this trend will be maintained in the future. Figure 3.3 shows a schematic view of a complete lighting system. To generate light that eventually reaches the eye, every system includes most or all the steps shown. In order to work in terms of power so that we can calculate efficiencies, the quantity in Figure 3.3 \( P_{vis} = \Phi_v \times 683 \) for each stage in there is a loss and the system efficiency is then

\[
\eta_{sys} = \eta_{dc} \times \eta_{cir} \times \eta_{rad} \times \eta_{VR} \times \eta_{vis} \times \eta_{eff} \quad (3.8)
\]

The various terms are defined in Figure 3.3. Each stage in this chain of light production needs to be examined to discover how system efficiency can be improved. Notice that Equation 3.8 applies equally well to a street lamp, a projector, a self-ballasted lamp, a fiber-optic illuminator and, if the conversion from mains to ac is omitted, to battery operated lighting system.

### 3.3.5 Color properties and color temperature of sources

Definitions of quantities mentioned below related to color are given in Chapter 8. A comprehensive discussion of color in lighting is also given by Coaton and Marsden [3, Chapter 3].

An important color property of any source is color appearance or chromaticity (specified by the chromaticity coordinates). The color appearance
of any source can be matched by mixture of three sources of different color appearance (for example, by red, green and blue sources, or by three spectral sources). The space of all possible colors is bounded by the spectral colors. For general illumination and for some optoelectronic applications such as projection, the preferred color of sources is “white”; the chromaticity of these sources is then close to that of a black body having a color temperature (see below) in the range from about 2800 (yellowish white) to about 8500 K (bluish white). Other sources, such as those used for signaling, usually have more saturated colors (that is, colors such as red, green, amber, etc.) that are close to the spectral colors. The specifications for these sources are closely controlled [9].

Color temperature is defined only for those sources having a color appearance close to that of a black body. The quantity most often used is the correlated color temperature (CCT) [10,11], defined in Chapter 8. A few examples help to set a scale. The glowing embers of a fire have a CCT in the region of 1000 K whilst a candle flame has a CCT of about 2000 K. Incandescent lamps, depending on type, have CCTs between 2400 and 3400 K. The CCT of the sun is about 6000 K. Discharge lamps for general illumination mostly have CCTs between 3000 and 6500 K. Xenon arcs and flash lamps have CCTs in excess of 6000 K.

Sources of a given chromaticity (that is, having the same color appearance) may have very different spectral distributions. A commonly observable example (at least in Europe) is that the color of an amber traffic signal and of the commonly used orange low-pressure sodium street lights are almost identical; the sodium lamp emits only at about 589 nm whereas the traffic signal is a filtered tungsten lamp that emits over a broad spectral range from the yellow through to the red. The value of $K$ (Equation 3.6) for light from the LP sodium lamp also greatly exceeds that for the traffic signal.

Not surprisingly a surface illuminated by these two sources appears to have very different colors. The color rendering capability of a light source is an important measure. For general task illumination colors need to appear “natural”; this means that surfaces such as skin, fabric, building materials, etc. should not appear distorted when compared with their appearance under natural light or incandescent light, which both have continuous spectra. Along with high efficacy or high luminance, this is a major requirement for commercial light sources. The measure of color rendering used is the CIE General Color Rendering Index (CRI) or $R_a$ (see Chapter 8) [12,23]. $R_a$ is computed from the color shifts shown by a series of colored surfaces when illuminated by the test illuminant as compared to their color when illuminated by natural and Planckian reference illuminants.

The better the color rendering, the lower the efficacy of the lamp. One might think that since the sources that give “perfect” color rendering have continuous spectra, then high quality lamps should too, and this is usually the case. However, simultaneous optimization of $K$ and $R_a$ at constant color temperature has shown a surprising result; both quantities are maximized if the light is emitted in narrow bands at 450, 540 and 610 nm. This feature of human vision, confirmed by experiment, has been exploited in the triphosphor fluorescent lamps that are now standard in all new installations. The lamps use narrow band phosphors that emit close to the critical wavelength. Similar techniques are now being used to optimize white light LEDs (Zukauskas et al. [4] give a useful review of optimization.)

The CIE CRI is defined so that tungsten and daylight sources have $R_a$ = 100. For general lighting in commercial premises requiring high-quality illumination, restaurants and homes, should be 80 or higher. Good quality sources for interior lighting such as triphosphor fluorescent lamps and HP ceramic metal halide (CMH) lamps have $R_a$ ≥ 80. Lower cost halophosphate fluorescent lamps have $R_a$ around 50–60, as do HP mercury lamps with phosphor coatings. High-pressure sodium (HPS) lamps used for street lighting have $R_a$ around 25.

Human vision is extremely sensitive to small differences in color [13] particularly in peripheral vision. This has proved to be a major challenge for lamp manufacturers, especially where lamps are used in large installations such as offices and stores. Not only should the initial spread in color be very small, but also the color shift during life must be very small otherwise when lamps are replaced it will be very obvious. Amongst the lamps for high-quality illumination triphosphor fluorescent and CMH are preeminent in this respect; such color differences as they have, are barely noticeable.
3.4 Incandescent lamps

3.4.6 Radiation from atoms and molecules in extended sources

In a discharge lamp, each elementary volume of plasma emits optical radiation. In a volume source an atom or molecule with an upper state of energy \( E_u \) (J) can make a transition to a lower state \( E_i \) (J) with a transition probability of \( A_{ul} \) (s\(^{-1}\)). The emitted wavelength \( \lambda \) (m) is then given by

\[
\frac{hc}{\lambda} = E_u - E_i(J) \tag{3.9}
\]

where \( h \) is Planck’s constant in Js\(^{-1}\) and \( c \) is the velocity of light in ms\(^{-1}\).

Since the emission is isotropic, the emission coefficient \( \varepsilon_{\lambda}(x) \) from a volume element at position \( x \) containing \( N_u \) atoms or molecules in the excited state is

\[
\varepsilon_{\lambda}(x) = \frac{10^{-9}}{4\pi} N_u(x) A_{ul} \times \frac{hc}{\lambda} P(\lambda) \text{ (Wm}^{-3}\text{sr}^{-1}\text{nm}^{-1}) \tag{3.10}
\]

\( P(\lambda) \) is the line shape function having an area normalized to unity. Do not confuse the emission coefficient \( \varepsilon_{\lambda} \) \([14]\) with the spectral emittance \( \varepsilon(\lambda, T, \theta) \) of a surface, which is a dimensionless quantity.

Suppose, we view a nonuniform extended source of depth \( D \). The spectral radiance along a line of sight for a spectral line at wavelength \( \lambda \) is

\[
L(\lambda) = 10^{-9} \times \int_0^D \varepsilon_{\lambda}(x) dx \text{ (Wm}^{-2}\text{sr}^{-1}\text{nm}^{-1}). \tag{3.11}
\]

This is only an approximation. When absorption is present the radiance does not depend linearly on atom density and is given by the radiation transport equation \([14]\). Examples of this important phenomenon are described in Sections 3.6.1 and 3.7.1.

3.4 INCANDESCENT LAMPS

3.4.1 Emission

Tungsten is the preeminent material for the manufacture of incandescent lamps. It has a melting point of 3680 K and it can be drawn into the fine wire necessary for making lamps. In normal household bulbs the filament is operated at a temperature in the region of 2800 K, depending on the type. The comparatively low temperature is chosen to limit evaporation and give an acceptable life. This section will concentrate on the higher temperature tungsten–halogen lamps that have many optical applications.

A substantial fraction of radiation from a tungsten filament is emitted between 750 nm and the glass or silica cut-off in the IR. Figure 3.4 shows that the fraction of power radiated in the region 750–2000 nm is approximately independent of the tungsten temperature, whilst the visible fraction (<750 nm) doubles for an increase of 500 K in temperature.

3.4.2 Tungsten–halogen lamps

Use of a halogen chemical transport cycle \([15]\) allows tungsten filaments to be operated at higher temperatures than in the standard household bulb. For lamps of similar wattage and life the filament can be operated 100 K higher in a halogen lamp compared with a conventional lamp \([16]\).

The halogen—usually a fraction of a \( \mu \text{mol cm}^{-3} \) of iodine or bromine—is added to the lamp before it is sealed. During operation of the lamp the halogen reacts with evaporated tungsten in the cooler regions. The tungsten halide thus produced is a
vapor that is transported by diffusion and convection to hotter regions, where it dissociates depositing tungsten and releasing halogen for further cleanup. The dissociation mainly takes place at a region of the filament lead. The net effect of the cycle is therefore to transport the tungsten from the wall to regions of the lamp that do not affect light output.

Because the lamp walls remain clean, the bulb can be made very small and strong. High pressures of inert gas of high molecular weight suppress evaporation. With smaller, stronger bulbs containing high pressures of Kr, or even Xe, the tungsten may be operated at temperatures of up to about 3500 K.

The higher the filament temperature the greater the rate of evaporation and the shorter the life of the lamp. Filaments operating at a color temperature of 3400 K (filament temperature ≈ 3330 K) will have a life of a few tens of hours. Life is also strongly dependent on operating voltage; manufacturers’ data should be consulted for information.

Tungsten–halogen lamps have the advantage over all other sources of having excellent stability. For best stability, lamps should be operated from a dc constant current supply with current controlled to 1 part in $10^4$—this is the technique used for operating calibration lamps. The current should be set to ensure that the voltage rating of the lamp is not exceeded. When lamp stability is at premium (as for example in standards of spectral irradiance) optical equipment suppliers select particularly stable lamps (Appendix—Oriel Instruments, Ealing).

### 3.4.3 Varieties of tungsten–halogen lamps

The development of tungsten–halogen lamps has resulted in thousands of new products being introduced. For optical applications the most important consideration is often the ability to focus the light into a tight beam. This is affected by the size of the filament, the tightness and evenness of winding of the coil, whether the coil is a flat or cylindrical, whether the coil is concentric with the bulb axis or normal to it, the quality and thickness of the bulb wall, and the type of glass used (hard glass can have better optical quality than fused silica). Examples are shown in Figure 3.5. High color temperature versions with powers in the range 25–1000 W or even greater are available. In some cases, these are made from silica that is doped to prevent emission of short wave UV. Consult manufacturers’ websites for “special” lamps designed for particular optical applications.

### 3.4.4 Lamps with integral reflectors

There is a wide range of tungsten–halogen lamps built into small reflectors. The reflectors may be aluminized, or have a dichroic (interference filter) coating allowing some IR radiation to escape from the rear of the reflector; this means that the beam is comparatively cool. They may also be fitted with cover glasses that reduce the already small amount of short wave UV that is emitted by fused silica tungsten–halogen lamps. Reflector diameters vary from 50 down to 35 mm. Versions are made with beam divergences from a few degrees up to 40°. Typically wattages vary from 12 to 75 W with a color temperature of about 3000 K.

These reflector lamps are used in large numbers for all sorts of commercial displays and accent lighting and therefore they are relatively inexpensive. In addition, all the major lamp manufacturers make special versions that are used in a number of optical applications such as overhead projection, microfilm and fiber optic illuminators. Figure 3.6 shows examples.
3.5 Discharge lamps with electrodes

3.4.5 Lamps with IR reflectors

Over 90% of the radiation from tungsten–halogen lamps is in the IR region and so is wasted. Many attempts have been made to return some of this radiation to the filament where it can be absorbed. Commercial success was eventually achieved by using multilayer interference filters deposited by LP CVD [16]. Less input power is needed to maintain the tungsten coil at the design temperature. The main benefit therefore is a saving in power for a given light output. The beam is cooler since there is less IR radiation emitted although optical quality is degraded slightly by the coating. The main benefit is an improvement of up to 40% in the efficiency of generation of visible light.

3.4.6 IR sources

Incandescent lamps using either tungsten or carbon emitters make use of the IR radiation in industrial heating processes (Appendix—Heraeus). The main benefit is a heat source that can be controlled precisely and has a much shorter response time than a conventional oven.

The Nernst source is an example of a ceramic emitter electrically heated to 2000 K, used as an IR illuminator in spectrophotometers. This makes use of selective emittance in the IR. More recent versions of similar devices are given in manufacturers’ data (see Appendix—Oriel Instruments). There are also low heat capacity carbon emitters that can be modulated at low frequencies (Appendix—Hereaus).

3.5 DISCHARGE LAMPS WITH ELECTRODES

One way is to group discharge lamps into LP (low-pressure) and HP (high-pressure) types. In LP discharges, the electrons make relatively few collisions per second with the gas atoms and so electron temperature $\gg$ gas temperature. In HP discharges, relatively frequent collisions between electrons and gas atoms ensure that both temperatures are approximately equal. The same physical processes occur in LP and HP discharges. Section 3.5.1 is concerned with the common features of both types. The electrode regions are described in Section 3.5.2. Later sections describe their unique features.

Another way to group discharges is by the manner of coupling to the power supply. Most discharges have electrodes in which the cathode is hot; electrons are released into the plasma by thermionic emission. The term arc is not uniquely defined, but it is often taken to mean a discharge in which the cathode emits thermionically—examples are all HP discharge lamps and hot cathode fluorescent lamps. In cold cathode lamps, the electrodes emit as a result of ion bombardment of the cathode surface. Other discharges (Section 3.9) are operated at high frequency using induction or microwave sources. Dielectric barrier discharges (DBDs) are transient and self-limiting with little or no emission of electrons from the cathode (Section 3.9.2).

3.5.1 Stable discharge operation of discharges with electrodes

To start a discharge, a high voltage must be applied to make the gas conducting, and (an electron) current from an external circuit must be passed from cathode to anode through the conducting gas. A by-product of causing the gas to conduct is the production of radiation. To demonstrate the main effects we will consider dc discharges although the majority of commercial lamps operate on ac (Section 3.8.3).

We will illustrate the main features of a dc discharge using the LP mercury–rare-gas discharge of the type used in fluorescent lamps as an example (Figure 3.7). Other lamps including HP lamps have...
50 Incandescent, discharge, and arc lamp sources

similar features, but the regions around the electrodes have dimensions that are usually too small to see. The bulk of discharge in Figure 3.7—the positive column (PC)—is a plasma, so there are equal number of electrons and ions per unit volume. Some discharges such as neon indicators or deuterium lamps used for producing UV are so small that the PC does not exist.

In the PC, electrons form a near Maxwellian distribution of energies. Once the discharge has been established, the applied electric field causes the electrons to drift towards the anode and the ions to drift towards the cathode; because their mobility is much greater than that of the ions and the current is carried mainly by the electrons. Therefore, current density is approximately

$$ j = n_e |e| \mu_e E \left( \text{A m}^{-2} \right) $$

(3.12)

where $E$ is the electric field, $\mu_e$ the electron mobility, $|e|$ the electron charge and $n_e$ the electron density.

The PC can be any length as long as sufficient open-circuit voltage is available from the supply (think of commercial display signs). A condition for stable operation is that the rate of loss of electrons by recombination with ions must be equal to the rate of gain caused by ionization. In LP discharges, most of the recombination occurs after the carriers have diffused to the wall; in HP discharges, particle densities are high enough for volume recombination to dominate.

The electric field in the column adjusts itself so that electrons are accelerated to a mean energy in the region of 0.5–1.5 eV corresponding to an electron temperature of about $T_e$ of 6000–18000 K. The electron energy distribution then contains enough high-energy electrons to ionize atoms, replacing the electrons lost by recombination. Figure 3.7 shows that the electric field in the PC is constant, so in a given gas, the longer the lamp the higher the voltage.

### 3.5.2 Electrode regions

Adjacent to the anode the voltage usually increases (Figure 3.7). This is a result of a space charge sheath. If there was no sheath then the anode would only collect the random current. Normally the anode area is too small; to collect the current required it charges positively to attract electrons.

The cathode is more complex [2, Chapter 4]. The conditions at the cathode surface have to adjust themselves so that each electron that leaves the cathode initiates events that cause the emission of at least one more electron from the cathode, otherwise the discharge will not be self-sustaining. Electrons emitted thermionically (hot cathode
3.6 Types of LP discharges

By far the most important type is the LP mercury rare-gas discharge used in fluorescent lamps and in UV sources for photochemical and photobiological purposes (Section 3.6.1). Other LP discharges not described here are LP sodium, used for street lighting of very high luminous efficiency, deuterium lamps used as UV illuminators and LP hollow cathode spectral sources for chemical analysis. There are also a wide variety of LP laser discharges.

### 3.6.1 Low-pressure mercury rare-gas discharges

LP mercury lamps contain a rare gas, usually argon, krypton or neon or mixtures of these, at a pressure of a few hundred pascal (a few torr). Mercury is added as a small drop of liquid weighing a few milligrams, which collects at the coolest place in the lamp. At typical wall temperatures, the mercury evaporates from the liquid drop at the pressure of about 0.8 Pa (0.6 m Torr). Despite the relatively low number density of the mercury atoms they dominate the properties of the discharge. The fluorescent lamp discharge is a highly efficient emitter of UV in the mercury resonance lines at 254 and 185 nm (>70%).

Phosphors are used to convert UV to visible radiation [3, Chapter 7]. Lamp phosphors are ionic materials doped with activators that absorb at short wavelengths and then reemit at longer wavelengths. The energy deficit in this *Stokes’* shift is converted into lattice vibrations. In fluorescent lamps used in lighting the conversion loss is typically 50%. There is a very large range of phosphors [3, Chapter 7], and fluorescent lamps giving white light of many different CCTs and other color properties are available. Particularly important are the ionic rare-earth based phosphors as these emit at the wavelengths that combine high efficacy and color rendering index (*R*ₐ—Section 3.3.5). The principle ones are noted in Table 3.1; notice how close the peaks are to the 450, 540 and 610 nm wavelengths that optimize color rendering and efficacy.

### Table 3.1 Phosphors commonly used in fluorescent lamps

<table>
<thead>
<tr>
<th>Name</th>
<th>Formula</th>
<th>Wavelength of peak output (nm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>YEO</td>
<td>Y₂O₃:Eu³⁺</td>
<td>611</td>
</tr>
<tr>
<td>CAT</td>
<td>Ce₀.₆₅ Tb₀.₃₅ MgAl₁₁ O₁₉</td>
<td>543</td>
</tr>
<tr>
<td>LAP</td>
<td>LaPO₄:Ce³⁺, Tb³⁺</td>
<td>544</td>
</tr>
<tr>
<td>CBT</td>
<td>GdMgB₂O₁₀:Eu²⁺</td>
<td>545</td>
</tr>
<tr>
<td>BAM</td>
<td>BaMg₂Al₆O₁₆:Eu²⁺</td>
<td>450</td>
</tr>
<tr>
<td>Halophosphate</td>
<td>Ca₅(PO₄)₃(F, Cl):Sb³⁺, Mn²⁺</td>
<td>Broad bands</td>
</tr>
</tbody>
</table>
The notation in Table 3.1 is the chemical composition of host lattice:activator. The activator is an ion added deliberately at relatively small concentrations to absorb UV and emit visible light. In some cases, the host lattice has this same function. Quantum efficiencies are close to unity.

One very important benefit of rare-earth phosphors is their resistance to degradation by mercury discharges at high power loadings. It is this property that made possible the development of compact fluorescent lamps. The disadvantage is the high cost of rare-earth phosphors compared with the halophosphates that they have largely replaced. The complexity of the materials in Table 3.1 is such that phosphor research is still largely empirical, so the existence of each of these phosphors represent many man-years of painstaking research.

The mercury vapor pressure is a dominant factor in controlling the amount of radiation emitted and the efficiency with which it is generated. When a mercury atom is excited near the center of lamp, the emitted photon is at exactly the correct energy to be absorbed by a ground state atom nearby. The photon is absorbed and reabsorbed many times before it finally reaches the wall in a random walk. When the mercury pressure is high, there are so many steps in the random walk that the chance of losing the excitation energy nonradiatively in a collision increases. When the mercury vapor pressure is low the initial excitation energy can escape in a small number of steps, but then the fraction of collisions that lead to excited mercury atoms is low. (The related process in HP lamps is described in Section 3.7.1.)

This means there is a mercury vapor pressure at which the efficiency of generation of UV radiation is at a maximum. This optimum pressure is achieved by having a small amount of liquid mercury present at about 42°C. When using fluorescent lamps it is important to arrange for the fixture or unit holding the lamps to operate so that the mercury pressure is close to optimum. Lamps are designed to run close to optimum in commercial lighting fixtures. For other uses, such as backlighting some cooling may be necessary. Some types of multilimb compact fluorescent lamps are designed for operation in hot fixtures. In these, the mercury is dosed as a solid amalgam containing, for example, bismuth and indium. The vapor pressure of mercury above the amalgam is less than that above free mercury, but the use of an amalgam also substantially increases the ambient temperature range over which the mercury pressure is close to optimum [18].

### 3.6.2 Applications of LP mercury discharges

The fluorescent lamp discharge lends itself to many different formats [3, Chapter 7]. The most familiar are the long thin lamps used in ceiling lighting in nearly all commercial and industrial premises. There are also a wide variety of compact fluorescent (CFL) designed as a high efficiency replacement for incandescent lighting.

Other than illumination, important applications for fluorescent lamps are in office equipment (copiers, fax machines, etc.) and in the backlighting of displays. Cold cathode fluorescent lamps have a number of benefits: they can be small in diameter allowing screens to be very thin; at the low powers needed they can be efficient enough for the purpose; lives are long; they can be switched frequently; and low cost, efficient power supplies are readily incorporated in the end product. Hot cathode fluorescent lamps produce more light and can be used to backlight displays that are used in high ambient light levels such as ATM machines. Short wave radiation from hot cathode mercury rare-gas discharges is used in photochemical or photobiological processes; or it can be converted using a phosphor to UVA (as in “black light” sources) that show up fluorescence in materials.

### 3.7 HP DISCHARGES

There are many variants of HP discharges. Most of them are used for street lighting and interior illumination of stores and offices and other commercial premises, in which high luminous flux, high efficacy, good color quality and long life are at a premium. Lamps exist in single-ended (both connections at one end) and double-ended (one connection at each end) configurations to suit different applications. Many other types of HP discharges are used in which light must be projected and high brightness is needed. Some of the properties of HP discharges are described below. The two main classes of lamp are those that use volatile or gaseous elements, and those that use metal halides to introduce radiating species into the vapor.
### 3.7.1 General features of HP discharge lamps

We will illustrate the operation of HP discharges by using the HPS (high-pressure sodium) lamp as an example. An HPS lamp has electrodes inserted into a narrow arc tube made from translucent alumina, resistant to attack from sodium. As with many HP lamps the arc tube is contained within a glass outer bulb. These lamps are used as highly efficient (120 lm/lamp watt) long-lived (>20,000 h) street lights that give a pleasant golden light with CCT = 2000 K, albeit with rather poor color rendering properties (Ra = 25).

The dimensions of the tube are typically 7 mm internal diameter with 70 mm length between the electrode tips for 400 W rating, with dimensions decreasing for lower wattage lamps. They contain a small pressure of rare gas and a few milligram of sodium metal. On applying a voltage the rare gas breaks down. The resulting discharge heats and evaporates sodium until its pressure is about 1.4 × 10^4 Pa (100 Torr). A radial temperature profile develops in which the center temperature is about 4000 K and the wall temperature is about 1500 K. Most of the length of the discharge is a positive column uniform along the axial direction. Sodium lamps usually also contain about 10^5 Pa (760 Torr) of mercury vapor. This reduces thermal conduction and increases axis temperature, thus increasing spectral radiance.

The positive column is approximately in local thermodynamic equilibrium (LTE) [14]. This means that the properties are dependent on the local temperature in the plasma. The electron density is given by a version of the law of mass action called the Saha equation [19]

\[
\frac{n_e n_i}{n_a} = S(T) \ (\text{m}^{-3})
\]

(3.13)

where \(n_e\) and \(n_i\) are the electron and ion densities, \(n_a\) is the density of atoms (number per m^3) and

\[
S(T) = 4.83 \times 10^{21} \left( \frac{U_i/\epsilon_i}{T} \right)^{3/2} \times \exp(-E_i/kT) \ (\text{m}^{-3})
\]

(3.14)

where the \(U\) factors are partition functions for the ion and atom. \(S(T)\) depends strongly on temperature through the exponential factor, where \(E_i\) is the ionization potential (including corrections for high electron density) and \(k\) is Boltzmann’s constant. Since the hot gas is a plasma \(n_e = n_i\). The atom density \(n_a\) in an elementary volume at temperature \(T\) is given by the gas law so \(n_a = P/kT\) where \(P\) is the gas pressure. Table 3.2 shows values of \(n_e\) in sodium vapor at various temperatures. Since the current density is proportional to \(n_e\) it is clear that the current flow is mainly in the high temperature region.

The population \(n_u\) of an energy level of an atom (labeled \(u\)) is given by another LTE formula:

\[
n_u = \frac{g_u}{g_0} n_0 \exp \left( -\frac{E_u}{kT} \right) \ (\text{m}^{-3})
\]

(3.15)

where \(n_0\) is the density of atoms in the ground state, \(E_u\) is the energy (J) of the upper state of the atom, whilst \(g_0\) and \(g_u\) are the statistical weights of ground and upper states, respectively. The number of atoms excited to the upper state depends

<table>
<thead>
<tr>
<th>Plasma temperature (K)</th>
<th>2000</th>
<th>3000</th>
<th>4000</th>
<th>5000</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number density of sodium atoms</td>
<td>(4.8 \times 10^{23})</td>
<td>(3.2 \times 10^{23})</td>
<td>(2.4 \times 10^{23})</td>
<td>(1.9 \times 10^{23})</td>
</tr>
<tr>
<td>Fraction of sodium atoms excited to the states radiating at 589</td>
<td>(1.5 \times 10^{-5})</td>
<td>(8.8 \times 10^{-4})</td>
<td>(6.4 \times 10^{-3})</td>
<td>(2.3 \times 10^{-2})</td>
</tr>
<tr>
<td>Fraction of sodium atoms that are ionized</td>
<td>(7.3 \times 10^{-6})</td>
<td>(1.9 \times 10^{-3})</td>
<td>(3.5 \times 10^{-2})</td>
<td>(2.2 \times 10^{-1})</td>
</tr>
</tbody>
</table>

Note: Electron density is equal to ion density. There are two excited states at about 2.1 eV giving rise to the characteristic orange sodium D radiation. The ionization potential is 5.14 eV before correction is made for lowering of the value at high electron densities. The arc operates so that the electron density is sufficient to carry the current and the plasma temperature adjusts to make this so. For steady state sodium arcs this sets the maximum plasma temperature to about 4000 K. Calculated using Equations 3.13 and 3.15.
Incandescent, discharge, and arc lamp sources

exponentially on temperature. Because of the exponential Boltzmann factor in Equation 3.15, the fraction of atoms in the excited state \( u \) is very small even at the highest temperatures. Only in the hottest parts of the discharge are significant numbers of atoms excited; the resulting “corded” appearance is a characteristic feature of an LTE arc. When a HP discharge operates horizontally convection bows the bright part upwards—the origin of the term arc. The importance of Equations 3.13 and 3.15 is shown in Table 3.2.

Self-absorption dominates the spectrum of many HP discharge lamps and is especially dominant in HP sodium discharges. As Figure 3.8 shows there is no significant radiation at 589 nm, the wavelength at which sodium radiates at low pressures. In an HPS lamp, the sodium pressure is so high that photons from excited sodium atoms can only travel about \( 10^{-7} \) m at the line center before being absorbed by a ground state atom. However, there is a chance that very close collisions with other sodium atoms can perturb the radiating atom sufficiently so that it radiates at wavelengths far from the line center at 589 nm. The hot plasma can therefore be considered as storing excitation energy until the energy can escape from an atom having strongly perturbed energy levels. The higher the pressure, the further from the line center the wavelength has to be, before the light can escape (Figure 3.8). This behavior is called self-reversal and it has a dominating effect on the operation of many HP discharges [3, Section 5.6.3]. The cover of the book by de Groot and van Vliet [19] shows beautiful color photographs of the self-reversal of the sodium D lines at different pressures.

3.7.2 HP metal halide lamps

There are very few elements that have well-placed spectral lines and sufficiently high vapor pressures to be operated as HP discharges, the most important being mercury, sodium, sulfur, and the permanent gases (of which Xe is by far the most important).

There are perhaps 50 elements that have metal halides that are sufficiently volatile to be used in HP lamps. The principal ones are as follows:

- Na, I, In, Tl, Ga halides in which the metals have relatively few strong atomic lines.
- Sc, Fe, Dy (and other rare earth) halides in which the metals have many relatively weak visible lines so close together that the spectrum appears continuous at moderate resolution.
- Sn, Pb, and similar halides that form relatively stable monohalide molecules that emit a spectrum that appears continuous at low resolution.

The halide is usually the iodide, which has the least reactive chemistry.

Metal halide lamps are extensively used as efficient white light sources of good color quality for
3.7 HP discharges

3.7.3 Operating principles of metal halide lamps

Many metal halide lamps contain thallium iodide (TlI). TlI is considered as a simple case to illustrate how the light is produced in metal halide discharges. Figure 3.9 shows a schematic diagram of a HP TlI discharge. When the lamp is made, a few milligram of solid TlI and a rare gas for starting are added. Usually enough mercury is added to give a partial pressure of about 10^6 Pa (10 bar) to reduce thermal conduction and to adjust operating voltage (Section 3.7.1). When the lamp is operated, the rare gas discharge heats up the Hg and TlI causing them to evaporate. In higher temperature regions, the TlI dissociates into Tl and I atoms. At higher temperatures still the Tl is excited and emits intense green light of high efficiency that can be useful for underwater illumination. Finally, near the axis the Tl is ionized producing the electrons needed to carry the current. This progressive evaporation, dissociation, excitation and ionization occurs in all metal halide discharges.

With mixtures of halides the ratio of salts has to be chosen with due consideration to the chemistry of the liquids and vapors. For example, one of the first types of metal halide lamp used mixtures of indium, thallium and sodium iodides that emit blue, green and orange self-reversed spectral lines. Altering the proportions of these can provide white light discharges of different color temperatures and quite good efficacy (luminous efficiency). However, their color rendition is rather poor.

Metal halide arc tubes are generally shorter than HPS lamps (for which the length to diameter ratio is more constrained by requirements of optimization) and may even be close to spherical in shape. This has an effect on étendue and may make fixtures using these lamps more efficient at using the light.

There has been extensive research and development over the last 40 years that has produced mixed metal halide lamps with much improved color performance and efficacy. The halides used, their vapor pressures and their relative proportions all have a strong influence on the initial color properties and efficacy.

It is important that these properties stay constant through lives of 10,000 h or more. Reactions between the various components and the tube walls occur at different rates; all metal halide lamps show some color shift during life. Detailed R&D has improved the color stability of metal halide lamps in silica arc tubes so that it is acceptable in critical applications such as the lighting of stores and offices. A recent major improvement has been in the use of translucent alumina ceramic arc tubes for containing Na, Dy, Tl, HgI metal halide arcs.
Metal halide reactions with the envelope are much slower than with silica and this has provided a further major improvement in initial color uniformity and color stability through life.

Most metal halide lamps are used for illumination where the transparency of the arc tube is not an issue, but the scattering by the arc tube is a major disadvantage for projection. For projector and automotive head lamps, silica arc tubes are universally used. Because the axis temperature of metal halides is usually around 5500 K, the radiation of the gas close to the axis can be very high (see Figure 3.1).

### 3.7.4 Applications of HP discharge lamps

Table 3.3 gives information about HP lamps used for general illumination. It is intended to show the range of types available with some idea of the best characteristics to be expected. All types exist in more than one power rating, but the rating given is a fairly typical one for the indicated application and type. Generally efficacy increases as power rating increases [3, Appendix 1].

Most of the lamps that are in the table are arcs with positive columns of length of several centimeters that are stabilized by the tube wall. The CMH lamp is a short arc lamp in which the arc is mainly stabilized by the electrodes.

Manufacturers’ websites give many examples of applications other than for illumination. All the major lamps manufacturers make a variety of metal-halide and xenon short arc lamps for projection and related uses (Appendix). In short arc lamps (length < few millimeters) there are usually regions close to the electrodes that have particularly high arc temperature. This region of high arc temperature forms because the electrodes cool the arc, and the field close to the electrodes has to increase to maintain conduction; moreover the current density normally increases as the arc contracts toward the cathode hot spot. The combined increase in current density and field means that the power per unit volume of arc is greatest just adjacent to the electrode. Although this generally leads to a reduction of efficacy there may be an increase in luminance. Figure 3.10 shows examples of lamps that are used for a variety of projection and entertainment applications and other more specialized applications such as solar simulators.

Short arc metal-halide lamps can readily be integrated into prefocused parabolic or elliptical reflectors (Appendix—Welch Allyn, Ushio and others). Various beam divergences are available to suit different applications. A relatively recent development is shown in Figure 3.11 (Appendix—Philips, Osram and GE Lighting). This very high pressure (1.5×10^7 Pa or 150 bar) has extremely high luminance because of its extremely high arc temperature. The reason for the high arc temperature is that 130 W are dissipated in an arc of length hardly more than a millimeter. Spectral lines show extreme broadening and there is an intense continuum giving good color rendition. With an arc gap of only 1.2 mm the étendue is very small. The lamps are designed to be operated in a prefocused reflector and the whole assembly used in data or video projectors.

<table>
<thead>
<tr>
<th>Lamp type</th>
<th>Application examples</th>
<th>Power (W)</th>
<th>Initial (lm W⁻¹)</th>
<th>Life (10³ h)</th>
<th>CCT (K)</th>
<th>Rₚ</th>
</tr>
</thead>
<tbody>
<tr>
<td>HPS</td>
<td>Road lighting</td>
<td>400</td>
<td>125</td>
<td>30</td>
<td>2000</td>
<td>25</td>
</tr>
<tr>
<td>High CRI</td>
<td>Prestige town lighting</td>
<td>400</td>
<td>100</td>
<td>24</td>
<td>2200</td>
<td>60</td>
</tr>
<tr>
<td>HP mercury vapor+phosphor</td>
<td>Road lighting</td>
<td>400</td>
<td>60</td>
<td>24</td>
<td>3500</td>
<td>55</td>
</tr>
<tr>
<td>Metal halide</td>
<td>Prestige outdoor, stores</td>
<td>400</td>
<td>90</td>
<td>24</td>
<td>4000</td>
<td>70</td>
</tr>
<tr>
<td>CMH</td>
<td>Commercial interiors</td>
<td>100</td>
<td>90</td>
<td>12.5</td>
<td>3000</td>
<td>85</td>
</tr>
</tbody>
</table>
3.8 Electrical characteristics of discharges

3.8.1 Breakdown and starting in discharge lamps

The gas in the lamp must be converted from an excellent insulator into a good conductor with a resistance that can be as low as a few ohms. Figure 3.12 shows the voltage across the lamp as a function of current over a very wide range of currents. After breakdown, the current increases rapidly until finally it stabilizes at the value needed to satisfy the circuit equations. In order to start the lamp, the circuit must be able to provide a voltage in excess of the highest lamp voltage in this diagram.

In order to achieve breakdown some source of electrons is necessary. If not provided by other means, they result from ionization by cosmic rays or natural radioactivity in the materials of the lamp. In other cases reliable breakdown is aided by the addition of small amounts of radioactive materials such as Kr$^{85}$, or by photoemission from surfaces caused by a small external source of UV. In hot cathode fluorescent lamps the electrodes can be heated before the voltage is applied: at low temperatures the field-enhanced thermionic emission provides enough electrons [2]. In HP lamps, a third trigger electrode is often included adjacent to the main electrode. When the voltage is applied across this small gap, breakdown is assured; this gap then provides initial electrons for the main gap. A general rule is that the fewer the initial electrons the higher the starting voltage needs to be, and the longer the time lag before breakdown.

The majority of lamps operate with hot cathodes that emit thermionically. Once breakdown has been achieved the transition from abnormal glow to arc in Figure 3.12 must be achieved quickly and cleanly. Staying too long in the region of cold cathode operation, in which the electrons released...
Incandescent, discharge, and arc lamp sources

58

by ion bombardment can be very damaging and shortens lamp life, sometimes dramatically.

There are various schemes for starting fluorescent lamps [2] in which the electrodes are tungsten coils coated with electron emission mix. One of the most common is to use a starter switch in parallel with the lamp. The starter switch is wired so that when closed, current limited by the ballast is passed through both electrodes. Initially the starter switch is closed. This preheating process raises the temperature of the electrode to about 1000 K before the starter switch opens. On opening the switch the open circuit voltage plus the self-induced voltage across the inductance is applied across the lamp, causing breakdown. The main purpose of preheat ensures that thermionic emission occurs very soon after breakdown. The switch is usually a relatively inexpensive bimetallic type. The tolerances on starter switch operation are closely constrained according to the lamp type. Increasingly fluorescent lamps are operated from electronic ballasts. It is a relatively simple matter to include a precision electronic preheat circuit to enhance the lamp life. At the low cost end, preheat is not included; after breakdown the electrode is heated rapidly by ion bombardment until it reaches thermionic emitting temperatures. Up to this time secondary emission dominates but the ion bombardment heats the cathode; these so-called instant start lamps generally have shorter life than lamps that are preheated.

For HP lamps, preheating is not an option. Starting from cold when the pressure is around 10^4 Pa (0.1 bar), breakdown voltages in the region of some kilovolts are needed. After an HP lamp has stabilized the pressure may be many atmospheres; on turning off it will require some tens of kV to restart immediately. Various types of pulse ignitor are used. The speed of transition from the glow to the arc is sensitive to many factors related to electrode design, lamp fill, processing quality and the open circuit voltage available. Ballast and lamp designers work together to ensure that the glow to arc transition in Figure 3.12 occurs rapidly and cleanly to ensure long life.

3.8.2 Steady state electrical characteristics

For an ohmic conductor the number of carriers is independent of current, so changing the voltage simply changes the mean drift velocity of the carriers. As long as the temperature remains constant the current is proportional to voltage—Ohm’s law. If the temperature increases the carrier mobility decreases and the resistance increases. This is what happens in tungsten lamps in which the hot resistance can be 15 or more times higher than the cold resistance.

Discharges show strongly nonohmic behavior (Figure 3.12). A discharge is a current-controlled device; and the voltage between the terminals sets itself to maintain this current. An additional
impedance called a **ballast** is necessary to control the current. In response to increasing current hot cathode discharges respond by decreasing the voltage across their terminals. This is the so-called negative, or falling *V–I* characteristic. The rate of decrease of voltage with current is usually quite small. This corresponds to the arc region on the right hand end of Figure 3.12 where the lamp voltage is comparatively low.

Figure 3.12 shows what happens to voltage as the current is increased over many orders of magnitude. Increasing from low values of current the voltage decreases to a plateau region in which a glow is visible on the cathode. On increasing the current, the glow increases in area whilst the voltage remains constant, implying that the current density at the surface of the cathode is constant. This is called the **normal glow** regime. As the current is increased further, the glow finally covers all the cathode area and often the leads as well. At this point, the current density has to increase and the voltage across the terminals increases. This is called the **abnormal glow** region. In both the normal and abnormal regions, the major part of the lamp voltage is dropped across the CF. The resulting ion bombardment increases the cathode temperature and the cathode begins to emit thermionically and makes a transition to the arc regime, which has a low CF. The abnormal glow region has a positive resistance characteristic, but this is not stable unless there is a ballast in the circuit.

For a dc discharge a series resistance \( R \) is needed to stabilize the current \( I \). If the supply voltage is \( V_s \) and the lamp voltage \( V_T \) then

\[
V_T = V_s - IR(V) \tag{3.16}
\]

The right hand side of this equation is called the load line. Load lines for three resistances are shown in Figure 3.12. For the highest resistance the intersection point is in the normal glow region (typical of a neon indicator lamp). With the lowest resistance the intersection is in the arc region. The intermediate resistance has two intersections, the one at the lowest current is in the abnormal glow region. If the heating of the cathode is insufficient to cause a transition to an arc, then the lamp remains in the abnormal glow condition. In some cases when starting an arc the discharge sticks in the abnormal glow with a high cathode fall; the sputtering can then cause very rapid blackening of the walls and premature failure.

Despite what the manufacturers’ data sheets may say, Figure 3.12 suggests that there is no specific power at which a discharge must operate; adjusting lamp current by using ballast impedance and supply voltage means lamps may be operated at a wide range of powers—at least for a time. The consequences of operating at powers different from the rated power are usually a reduction in life; properties such as color temperature and color rendering and efficacy will also change. Nevertheless,
for specific applications this is an option that the user can consider.

The reason for needing a ballast is best explained by using an argument given by Waymouth [2, Chapter 2]. Figure 3.13a shows the falling arc characteristic (part of the right hand end of Figure 3.12). This characteristic is the locus of points for which $\frac{dn_e}{dt} = 0$. The further above this line the more the rate of ionization exceeds the loss, so the current increases, and this increases $\frac{dn_e}{dt}$, with the result that the current continues to increase. If the applied voltage is below the line the loss exceeds production, the current decreases and the discharge extinguishes. Figure 3.13b shows the effect of a series ballast resistor. The total circuit voltage $V_T + IR$ now intersects the supply voltage at a certain current. If a fluctuation causes the current to increase, then the total circuit voltage moves into a region where $\frac{dn_e}{dt} < 0$, thus immediately decreasing the current. If the current decreases, then the total circuit voltage decreases so that $\frac{dn_e}{dt} > 0$, thus increasing the current again.

3.8.3 AC operation

Resistive ballasts work satisfactorily, but are lossy. Commercial lamps operate from the ac mains supply using magnetic inductances as ballasts [3, Chapter 17]. Figure 3.14 shows the lamp voltage and current waveforms for a fluorescent lamp on a resistive ballast. At 50 Hz there is an appreciable restriking voltage after current zero. This extra voltage is needed to restore the electron density after it has decayed during the latter part of the previous cycle. If this restrike voltage exceeds the supply voltage the lamp will extinguish. The phase relationships in an inductive circuit mean that a large voltage is available at the time that the current reverses, so extinction is less likely. For stability on ac mains supplies with a series inductance, the rms lamp voltage should not exceed about half of the rms mains voltage.

Most lamps are now developed to operate from electronic power supplies. Although more expensive than magnetic ballasts, there are a number of benefits: in fluorescent lamps there is an improvement in efficiency of UV production because of reduction in electrode loss and an increase in PC efficiency; electronic circuits can also provide programmed start and run-up sequences that prolong lamp life; there is no perceptible 50 or 100 Hz flicker from lamps run from electronic circuits at high frequency; and there is no restrike peak. Figure 3.14 shows typical waveforms at 50 kHz. In the case of HP discharges, operation at high frequencies can cause acoustic resonances that result in gross movements or distortions of the arc [19]. The electronic option is then to operate the lamp from a commutated dc—a square wave with fast transition times at frequency in the region of 90–500 Hz. For HP discharges, the lack of flicker and the ability to control lamp power (and thus color) over life are important benefits.

The optical radiation from discharge sources fluctuates by a percent or two. Part of this is caused by small changes in the cathode termination resulting in arc movement. It has recently been found that modified square wave supply waveforms can reduce the movement of the arc termination on current reversal [20]. A great improvement in stability can also be achieved by measuring the light output and using it to adjust the power into the lamp (Appendix—Oriel Instruments, Light Intensity Control System). A similar device can also be used to control the already excellent stability of tungsten-halogen lamps.
3.9 OTHER METHODS OF EXCITATION OF DISCHARGES

3.9.1 Pulsed light sources

A number of lamps are designed for pulsed operation [21]. The obvious example is the xenon flash tube used for photography, laser pumping, as warning beacons and as a transient source for scientific studies. The duration of the flash is of the order of microseconds with repetition frequencies up to hundreds of hertz. Operation is by discharging a capacitor through the lamp. Peak currents may reach thousands of amperes and electrodes must be constructed accordingly. The effects of using pulsed or transient output are that electron temperature can reach substantially higher values than in steady state. The result is usually due to an enhancement of the short wavelength radiation and an increase in peak radiance.

3.9.2 Dielectric barrier discharges

A form of transient discharge, DBDs have been used for large-scale industrial processes such as ozone generation for water purification and for generating far UV radiation for photochemical processes. DBDs may be operated in the pressure range from about $10^2$ to $10^5$ Pa [22].

Recently a DBD light source, the Osram Planon lamp has been developed. This provides a very uniformly lit tile-shaped area of reasonable luminous efficiency. At present, lamps are made in square format with diagonals up to 540 mm having a uniform luminance of $>6000 \text{ cd m}^{-2}$ (Appendix—Osram) [23].

The operating principle of a DBD is as follows. High voltage pulses (of some kilovolts) are applied between two electrodes, at least one of which is covered with an insulator of high breakdown strength such as glass. On applying a high voltage pulse, electrons are accelerated towards the anode and form an avalanche that breaches the gap. Electrons arriving at the anode charge up the surface, thus reducing and finally reversing the electric field. Electron current flows first from cathode to anode and then, when the anode charges up, from anode to cathode. The discharge lasts for a time ~μs. During the off period the ionization decays, providing the starting conditions for the next pulse. The discharge therefore comprises a series of microdischarges with lateral extent approximately equal to the electrode spacing. Microdischarges occur every time the pulse is turned on. DBDs have extremely non-Maxwellian energy distributions in which there are many high energy electrons. Because of this the excitation of resonance states of rare-gas atoms and molecules is favored, leading to high efficiency of UV production.

The Planon lamp is formed from two glass plates. On the lower plate, a metal cathode interlaced with a metal anode structure is deposited. Both electrodes are coated with glass to form the barrier layers. This form of electrode structure results in very uniform illumination. The lamp is operated from an electronic power supply designed to produce the optimized pulse sequence that is necessary for high efficacy. The two plates are held apart by spacers and the whole structure is sealed and filled with Xe at about $1.4 \times 10^4$ (100 Torr). Xe forms an excimer Xe$_2^*$ (excited dimer) that radiates efficiently in the vacuum UV at about 172 nm. Phosphor on the inner walls converts the UV to visible radiation. The use of Xe means that the output is almost independent of the lamp temperature so the lamp works just as well outside in cold weather as it does in the confines of office equipment.

The main applications are in displays and office equipment applications where a uniform and high luminance is a requirement. Cylindrical lamps based on the same technology are used in multifunction copiers.

3.9.3 Excitation by induction and by microwaves

In the last decade, a number of inductively coupled lamps have become available commercially from the major lamp manufacturers [3, Chapter 11]. All are variants on the fluorescent lamp discharge. Figure 3.15 shows a particularly compact example. The coil in the center is driven at a frequency of about 2.6 MHz. The rate of change of magnetic flux induces a voltage in the azimuthal direction. This causes a current to flow in a torus surrounding the coil. The ballasting is the result of the internal impedance of the supply. Benefits are long life and compactness. Other versions are Philips QL which, with a life of 100,000 h, is designed for use in inaccessible fixtures. Typically these will be high-bay fixtures with lumen packages between 2800 and 9600 lm. The Osram Endura or Icetron lamp which
has a stretched torus configuration has higher efficiency and packages of 8000–12,000 lm and a rated life of 80,000 h.

Microwaves can also be used to excite discharges. Fusion Lighting has pioneered a HP sulfur discharge in which the radiation is emitted by S₂ molecules. The light is white with a CCT in the region of 6000 K and the efficiency of generation can be up to 170 lm/microwave watt—higher than any other white light source. The overall efficiency is reduced because of the relatively poor efficiency of generation of microwave power. Light output levels are very high so the source is used in lighting large buildings. The very high radiance of such sources means that optical means can be used to distribute the light efficiently around buildings.

**Figure 3.15** Inductively coupled discharge lamp (GE Genura). The schematic diagram on the left shows that the plasma is a toroid with inductance \( L_a \) and resistance \( R_a \) that acts as a secondary to the excitation coil. The primary of the circuit is an impedance \( R_1 + j\omega L_1 \) that includes the effects of the plasma impedance, which depends on the power dissipated in the plasma.

3.10 LEDs FROM THE PERSPECTIVE OF CONVENTIONAL LIGHTING

Chapter 10 gives detailed information about LEDs. The generation of light by conventional lamps is limited by the black body radiance at the electron temperature. The reason is that the electron motion is randomized. In an LED, the motion of the carriers into the recombination region is far from random; the maximum radiance is therefore not limited by the Planck distribution. Already trichromatic LED [24] assemblies reaching 100 lm W⁻¹ have been made. (It is not clear in [4], and in many other LED publications, if this is the flux per lamp watt or flux per wall plug watt.) Note that these lamps are colorimetrically similar to the triphosphor lamps mentioned in Section 3.3.5. A further advantage of LEDs for many applications is their low étendue (Section 3.3.3) which means that the light can be directed more efficiently to where it is needed.

LEDs have already made a substantial impact on the conventional lamp manufacturing businesses. There are a number of applications where LEDs are far superior to conventional lamps. The obvious example is traffic signals. The LEDs generate the colored light only at the wavelengths needed, as compared with the filtered tungsten lamps used until recently. The required signal luminance values can therefore be met at much lower power consumption. It has been estimated that if all traffic lights in the USA were converted to LEDs, electricity consumption would be reduced by 0.4 GW [4]. Moreover, the LEDs are particularly well suited to withstand the vibration they experience as a result of heavy traffic and wind, and they may be frequently switched without damage. But the main advantage is their long life, which dramatically reduces maintenance costs compared with conventional traffic lights. A further advantage is that catastrophic failure of an LED does not have to cause the complete failure of the traffic signal, so safety is improved. All these advantages pay for the extra cost of the LED systems, so we can expect to see a complete takeover of the signal lamp market in land, sea and air transport signs.

White light LEDs are now finding many applications in decorative, aesthetic and artistic lighting where their properties are stimulating designers to produce interesting new ways of using light. There are also other niche markets such as highly localized task lighting where LEDs will make inroads, such as car interiors, desk, stairwell, path lighting, etc. In any application requiring relatively low flux levels, conventional lamps can now probably be replaced by LEDs as long as the substitution is not too expensive.

The ultimate target for the LED industry must be the replacement of the huge numbers of fluorescent lamps for lighting offices and other commercial premises. Much is made of the possibility of exceeding the efficiency of present day fluorescent lamps. The future is far from clear however: the issue is the low luminous flux from LEDs. No single LED approaches the level of flux required,
so any competitive installation will require many LEDs to produce the required flux levels, although the étendue advantage may reduce the flux levels required for LED installations. But the hard fact is that the cost of making an LED light source increases approximately linearly with flux, whereas the cost of making conventional lamps is very weakly dependent on flux. It is expected that manufacturing costs in both industries will be driven down by competition to comparable levels. If so, replacement of fluorescent lamps by LED equivalents on a global scale may come down to the cost of the materials used to make the lamps. In the case of fluorescent lamps, most of the material cost is in the rare-earth phosphors. Can the semiconductors used in making LEDs ever approach those costs per lumen? Will the potentially low cost organic light emitters ever be able to meet the flux requirements of commercial lighting? We shall see.

APPENDIX 3A

Selected manufacturers and suppliers of lamps

The manufacturers on this list give particularly helpful data in catalogues and/or websites for lamps with actual or potential electro-optical applications.

<table>
<thead>
<tr>
<th>Manufacturer</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cathodeon</td>
<td>Lamps for scientific instruments</td>
</tr>
<tr>
<td>Ealing Electro-Optics</td>
<td>Lamp units for integration into optical systems</td>
</tr>
<tr>
<td>Fusion Lighting</td>
<td>Microwave discharge lamps</td>
</tr>
<tr>
<td>GE Lighting</td>
<td>Full range of lamps for illumination and special purposes</td>
</tr>
<tr>
<td>Harrison Electrical</td>
<td>Cold cathode fluorescent</td>
</tr>
<tr>
<td>Heraeus Noblelight</td>
<td>Special lamps mainly for industrial and scientific processes</td>
</tr>
<tr>
<td>Iwasaki</td>
<td>Full range of lamps for illumination and special purposes</td>
</tr>
<tr>
<td>Osram</td>
<td>Full range of lamps for illumination and special purposes</td>
</tr>
<tr>
<td>Oriel Instruments</td>
<td>Lamp units for integration into optical systems, spectra of lamps</td>
</tr>
<tr>
<td>Philips Lighting</td>
<td>Full range of lamps for illumination and special purposes</td>
</tr>
<tr>
<td>Stanley</td>
<td>Cold cathode fluorescent</td>
</tr>
<tr>
<td>Toshiba Lighting</td>
<td>Full range of lamps for illumination and special purposes</td>
</tr>
<tr>
<td>Ushio</td>
<td>Wide range lamps for audio-visual, entertainment, photographic, scientific/media and industrial processes</td>
</tr>
<tr>
<td>Welch Allyn</td>
<td>Lamps for special applications</td>
</tr>
</tbody>
</table>

REFERENCES


