1

The need for compact blue-green lasers

1.1 A SHORT HISTORICAL OVERVIEW

For years after its invention in 1961, the laser was described as a remarkable tool in search of an application. However, by the late 1970s and early 1980s, a variety of applications had emerged that were limited in their implementation by lack of a suitable laser. The common thread running through these applications was the need for a powerful, compact, rugged, inexpensive source of light in the blue-green portion of the spectrum. The details varied greatly, depending on the application: some required tunability, some a fixed wavelength; some required a miniscule amount of optical power, others a great deal; some required continuous-wave (cw) oscillation, others rapid modulation.

In many of these applications, gas lasers – such as argon-ion or helium-cadmium lasers – were initially used to provide blue-green light, and in some cases were incorporated into commercial products; however, they could not satisfy the requirements of every application. The lasing wavelengths available from these lasers are fixed by the atomic transitions of the gas species, and some applications required a laser wavelength that is simply not available from an argon-ion or helium–cadmium laser. Other applications required a degree of tunability that is unavailable from a gas laser. In many of them, the limited lifetime, mechanical fragility, and relatively large size of gas lasers was a problem.

At about the same time, new options for generation of blue-green radiation began to appear, due to developments in other areas of laser science and technology. The development of highly efficient, high-power semiconductor diode lasers at wavelengths around 810 nm opened up the possibility of diode-pumping solid-state lasers, such as those based on neodymium-doped crystals and glasses. New and improved nonlinear materials made it practical to apply second-harmonic generation to the infrared outputs of these diode-pumped solid-state lasers to generate wavelengths in the blue-green regions of the spectrum. Demonstrations in

1986 of compact green sources based on intracavity frequency doubling of diode-pumped neodymium lasers by researchers at Spectra-Physics and Stanford University sparked tremendous interest in sources based on this approach. This interest has led to commercially-available diode-pumped green sources with powers of several watts and, more recently, blue sources with powers of several milliwatts.

Rather than pump a neodymium laser, why not simply use nonlinear optics to frequency double the output of an infrared semiconductor laser directly? The reason has been, until fairly recently, that high-power semiconductor diode lasers have had rather broad spectral distributions and rather poor spatial beam quality. While these characteristics did not prevent the use of these diode lasers as pumps for solid-state lasers, they did inhibit their use for direct nonlinear frequency conversion, in which the spectral and spatial mode properties of the infrared source are much more critical. As the spatial and spectral mode properties of high-power semiconductor diode lasers have improved, however, the same techniques of nonlinear frequency conversion have been applied to direct frequency-doubling of these devices, and efficient blue and green sources have been demonstrated. In some cases, resonant enhancement and guided-wave geometries have been used to increase the efficiencies of these nonlinear interactions.

An alternative approach to blue-green light generation using infrared sources is the so-called "upconversion laser". In a standard laser, energy conservation requires that the energy of an absorbed pump photon be greater than the energy of an emitted laser photon; hence the pump wavelength must be shorter than the lasing wavelength. In upconversion lasers, the energy from two or more pump photons is combined to excite the lasing transition; thus the pump wavelength can be longer than the lasing wavelength, so that, for example, infrared light can be used to directly pump a green laser. Although upconversion lasing was demonstrated in 1971 by Johnson and Guggenheim (1971), the field remained largely dormant for several years because flashlamp pumping of such lasers was inefficient. Experiments conducted at IBM in 1986 which demonstrated efficient laser pumping of upconversion lasers revived interest in the field. These initial experiments used bulk rare-earth-doped crystals and had to be performed at cryogenic temperatures, but they demonstrated the feasibility of these devices, including the fact that they could be efficiently pumped with laser diodes. Later, efficient room-temperature operation was achieved using optical fibers doped with rare-earth elements.

Perhaps the most direct and attractive way to generate blue and green light is to use a semiconductor diode laser. Semiconductor laser devices are efficient, small, robust, rugged, and powerful. However, in order to generate blue-green radiation, semiconductors with bandgaps of $\sim$3 eV must be used. Suitable materials systems include II–VI semiconductors such as ZnS and ZnSe, and wide-gap III–V materials such as GaN. The growth of thin films of these semiconductors suitable

for device fabrication has proven to be an extremely difficult challenge. However, breakthroughs in the growth of appropriately-doped films in both material systems has allowed the demonstration of light-emitting diodes (LEDs) and, more recently, lasers in both material systems. However, despite rapid progress, demonstration of continuous-wave (cw) operation at room temperature with powers and lifetimes comparable to infrared semiconductor lasers has not yet been achieved, and more development is required before these lasers can be used in the applications cited above.

1.2 APPLICATIONS FOR COMPACT BLUE-GREEN LASERS

One of the factors that has made the field of compact blue-green lasers interesting and vibrant is its diversity in both the variety of technical approaches used to produce them and the wide range of applications for which they have been sought. The specialized topical meetings that sprang up in the early 1990s in response to the intense interest and activity in this field (such as the Optical Society of America's Topical Meeting on Compact Blue-Green Lasers, held in 1992, 1993, and 1994) brought together researchers from such disparate fields as submarine communications and DNA sequencing. In this section, we review some of the principal applications for which blue-green lasers have been sought, and the requirements placed on the lasers by these uses.

1.2.1 Optical data storage

The terms "optical data storage" and "optical recording" have been used to refer to a variety of different approaches for recording and retrieving information using optical methods, including those based on such exotic phenomena as persistent spectral hole burning (Lenth *et al.*, 1986). However, these terms usually refer to somewhat more mundane systems that read data from (and, in some cases, write data to) spinning disks in a fashion analogous to magnetic disk drives (Figure 1.1).

In optical data storage systems, a bit is stored on the disk by altering some physical characteristic of the disk in a tiny spot. This alteration can be done once, as in the case of read-only disks (such as audio CDs and CD-ROMs), or it can be done repeatedly, as in the case of rewritable disks (such as those based on magneto-optic or phase-change media). To read back the information stored on an optical disk, a focused laser beam is scanned over these spots and the light reflected from the disk is detected. The physical characteristic that was altered to record a bit must produce a corresponding change in some optical property of the reflected beam. In audio CDs and CD-ROMs, data are impressed upon a plastic disk in the form of tiny pits stamped into the disk by the manufacturer. The depth of these pits is one-fourth
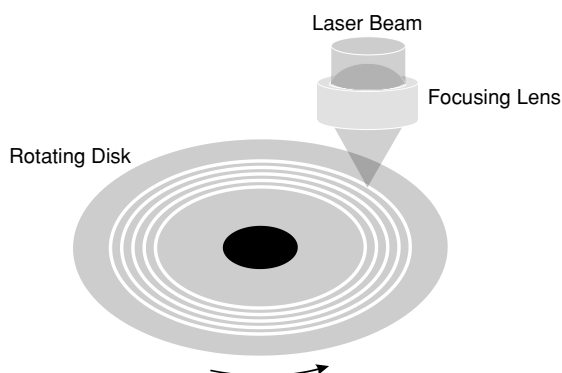
Figure 1.1: Optical data storage system.

the laser wavelength, so that when the beam is scanned over the pit, the portion reflected from the bottom of the pit travels an additional half-wavelength compared with the light reflected from the surface of the disk and is therefore 180° out-of-phase with it; thus, the amplitude of the reflected beam is diminished due to destructive interference. In "magneto-optic" media, data are recorded by using the laser beam as a heater: the focused laser spot heats the magnetic material above the Curie temperature, and the presence of an applied magnetic field causes the magnetization of the medium to reverse in the heated region. When the heating is removed and the material cools below the Curie temperature, that reversed magnetization is "frozen in". The data can be read back by exploiting the fact that the polarization of light reflected from the disk in these materials depends on the orientation of the magnetic domain (the "polar Kerr effect"). In "phase-change" material, data are recorded by using the focused laser beam to melt the material locally and induce a phase transition from what was originally a crystalline structure to an amorphous one. Data are read back by exploiting the fact that the amorphous state of the material has a different reflectivity than the crystalline state.

In order to write a small mark and be able to read it back accurately, the laser beam must be focused to as small a spot as possible. A gaussian beam can be focused by a lens to a diffraction-limited spot with a diameter d of

$$d \simeq \frac{\lambda}{NA}$$

where λ is the wavelength and *NA* is the numerical aperture of the lens. Therefore, one way to achieve a smaller spot size is to reduce the wavelength. Halving the wavelength from that of a GaAlAs laser diode at 860 nm to that of a blue laser at 430 nm would cut the spot size in half, and could quadruple the storage density. In addition, for a given rotation rate of the disk, the data rate could be increased by a

factor of 2, since the marks can be placed twice as close together. An additional motivation to pursue development of blue-green lasers for magneto-optic storage was the discovery of garnet-based recording materials that exhibit better performance in the blue-green regions of the spectrum than do their counterparts designed for use in the infrared (Eppler and Kryder, 1995).

Using a blue-green laser in an optical storage system places severe demands upon its performance (Kozlovsky, 1995). In the magneto-optic approach, the power required is comparable with that demanded of the infrared diode lasers used for optical storage ($\sim$40 mW) (Asthana, 1994). This may seem counterintuitive – one might expect that since the beam is focused to a smaller spot, less power would be required to produce the same temperature increase for writing. This statement is true as far as it goes; however, when reading data back with a blue beam, there are fewer photons per milliwatt than would be present in an infrared beam, which leads to increased noise. In order to obtain an adequate signal-to-noise ratio, the recording medium must be de-sensitized so that a higher readback power can be used without erasing the data. Thus, something like 2–6 mW is required for reading and 40–50 mW are required for writing. For focusing to a small spot, the wavefront aberration of the blue beam must be less than 0.05 wavelengths. The noise of the blue beam must be low: <-110 dBc (decibels below carrier) for magneto-optic storage, where differential detection is used, and <-135 dBc for phase change and CD-ROM, where single-ended detection is used. The laser must have a long lifetime, ideally as long as the lifetime of the drive itself (perhaps 100 000 hours mean-time-between-failures). Finally, the laser must be inexpensive.

1.2.2 Reprographics

Reprographic applications use lasers in a fashion similiar to optical data storage – the laser is focused to a small spot and used to make a mark on some medium. Here, however, the medium is the photoconductor of a laser printer, or photographic film or paper. Except in certain specialized applications (for example, writing on microfilm), reprographics does not require as small a spot size as optical data storage. A laser printer with 2400 dpi resolution requires that the laser beam be focused to only a 10 μm spot, a size that can be achieved easily using a red or near-infrared diode laser. However, this 10 μm spot size must be maintained as the beam is scanned rapidly over a page several centimeters wide. Decreasing the wavelength for a particular spot size relaxes the design requirements of the optical system by reducing the numerical aperture required to form a spot of the desired size and by increasing the depth-of-field.

In color reprographics, lasers can be used to expose photographic paper or film (Owens, 1992). The considerations just described for laser printers also apply here.

In addition, the wavelengths of the lasers must be chosen to provide correct exposure for existing photographic media. For photographic films, wavelengths of 430 nm, 550 nm, and 650 nm (blue, green, red) are desired. For photographic papers, wavelengths of 470 nm, 550 nm, and 700 nm are preferred. Powers of a few milliwatts are needed, along with good beam quality, low noise, and high stability. The ability to directly modulate the laser at frequencies up to 50 MHz is desirable.

1.2.3 Color displays

Blue-green lasers have also been sought for use in color displays. At present, the most popular type of color display device is the cathode ray tube (CRT) used in computer monitors and color televisions. In CRTs, colors are synthesized through the superposition of three primary colors – red, green, and blue – generated by an electron beam striking one of three corresponding phosphors. The combination of these red, green, and blue emissions in various proportions creates the other colors visible on the screen. A similar approach has been proposed for laser-based displays, in which three separate lasers would provide red, green, and blue primary colors that could be combined to project full-color images on a large screen (Figure 1.2). Each laser could be raster-scanned across the screen, or could remain stationary and be used to illuminate an "image gate", such as motion picture film or a spatial light modulator containing the image to be projected.

Lasers are attractive light sources for display applications because of their high brightness and complete color saturation. The brightness of a laser (power emitted per unit area per unit solid angle) can be very high due to the directionality of
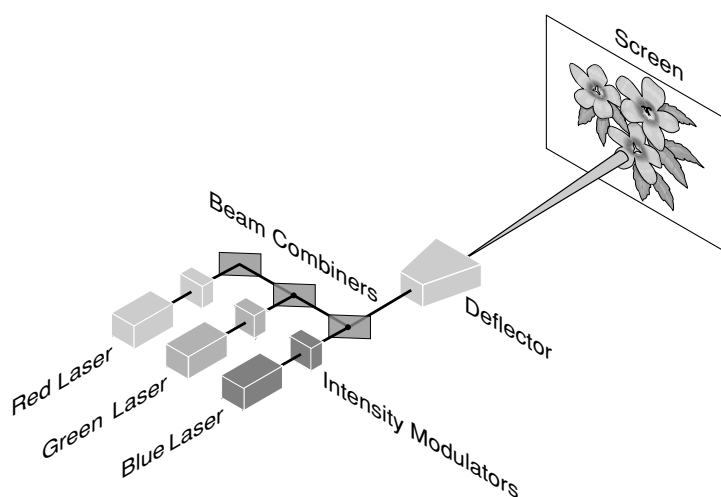

Figure 1.2: Laser-based color projection display.

the beam. This high brightness leads to high efficiency for a laser-based projector, since most of the generated optical power can be directed by appropriate optics to illuminate the screen or image gate. In contrast, a conventional motion picture projector uses an incandescent bulb that emits light into a 4π-steradian solid angle, most of which never reaches the screen. The ability of a laser to concentrate the emitted light into a confined solid angle provides an efficiency advantage over competing technologies (Glenn and Dixon, 1993).

Another advantage of laser-based displays is improved color saturation. In conventional CRT displays, the light emitted by the phosphor is not spectrally pure; the spectral bandwidth of the emission may be several nanometers. In the language of color theory, the red, green, and blue colors emitted by these phosphors are not "fully saturated": that is, the primary colors are not the "bluest blue", "greenest green", or "reddest red" that the eye can perceive, but appear somewhat washed out by the addition of white. As a result, a CRT cannot reproduce the entire range of colors perceptible to human vision, and in particular, cannot produce fully saturated colors. The range of colors that can be produced by addition of primaries can be depicted by the "CIE chromaticity diagram" (Figure 1.3). In this diagram, fully saturated colors (monochromatic light waves of a specified wavelength) correspond to points around the periphery. White corresponds to a point in the interior of the
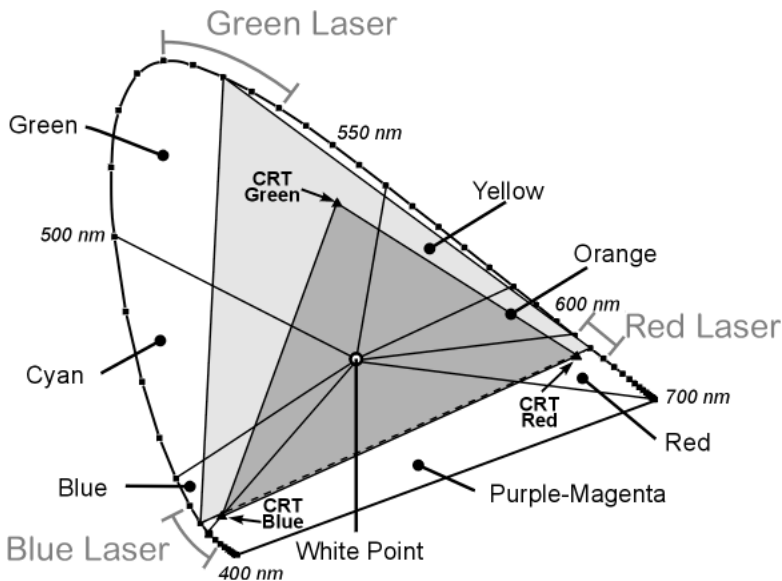

Figure 1.3: CIE diagram showing the color space spanned by CRT phosphors (dark shading) and the color space which could be spanned in a color display using monochromatic red, green, and blue lasers to generate the primary colors (lighter shading). The primary colors for each system fall at the corners of the triangles, as indicated.

diagram. If one draws a line from the "white point" out to a particular color on the periphery, the points along that line represent various saturation levels of the same color; for example, fully saturated green corresponds to a point on the periphery, and points along the line correspond to increasingly paler shades of green as one moves toward the white point. If one plots the points corresponding to three primary colors on such a diagram, the range of colors that can be synthesized by combining these primaries corresponds to the interior of the triangle connecting the three primary points. Figure 1.3 shows the points corresponding to the primary colors of a standard color CRT monitor. Although the red CRT phosphor is nearly saturated, the blue and green phosphors are considerably less so. Thus, while a CRT monitor can produce well-saturated reds, it is difficult to achieve well-saturated blues and greens. A laser-based color display produces primary colors that are fully saturated (that is, spectrally-pure monochromatic waves); thus, the range of colors that can be produced is greater and the colors themselves are richer than in a CRT. In order for the primary colors to appear to human vision as true blues, greens, and reds, they must fall within the wavelength ranges depicted in Figure 1.3: 605 nm $\pm$ 5 nm for red, 530 nm $\pm$ 10 nm for green, and 470 nm $\pm$ 10 nm for blue (Glenn and Dixon, 1993). The power required varies depending upon the size of the screen, but ranges from approximately 1 W per color for a 10-ft $\times$ 16-ft screen to 20 mW per color for a 16-in CRT-like display (Valley and Ansely, 1997).

1.2.4 Submarine communications

Communication with a submerged vessel has been another important application driving the development of blue-green lasers. Naval forces would like to be able to send messages to their submarines without requiring them to rise from their operating depth and risk detection by an enemy (Figure 1.4). Ideally, such a communications system would be able to send a signal through seawater to a great depth and to transmit information at a rapid rate. Electromagnetic radiation can penetrate seawater to a significant depth only at extremely low frequencies (ELFs) ($\lesssim 100$ Hz) or in the blue-green portion of the optical spectrum, where minimum attenuation (the "Jerlov Minimum") occurs for wavelengths between 400 nm and 500 nm (Figure 1.5). Although ELF systems have been built and used to send messages to submarines, systems using ELFs also have extremely low data rates, and in practice, only extremely short messages can be sent. Transmitting with blue-green wavelengths could make it possible to send messages to great depth with much higher data rate than with ELF. However simple this may sound in principle, the development of such a system has presented such great technical challenges that it has been described as "the most complex communications system known to man" (Painter, 1989).
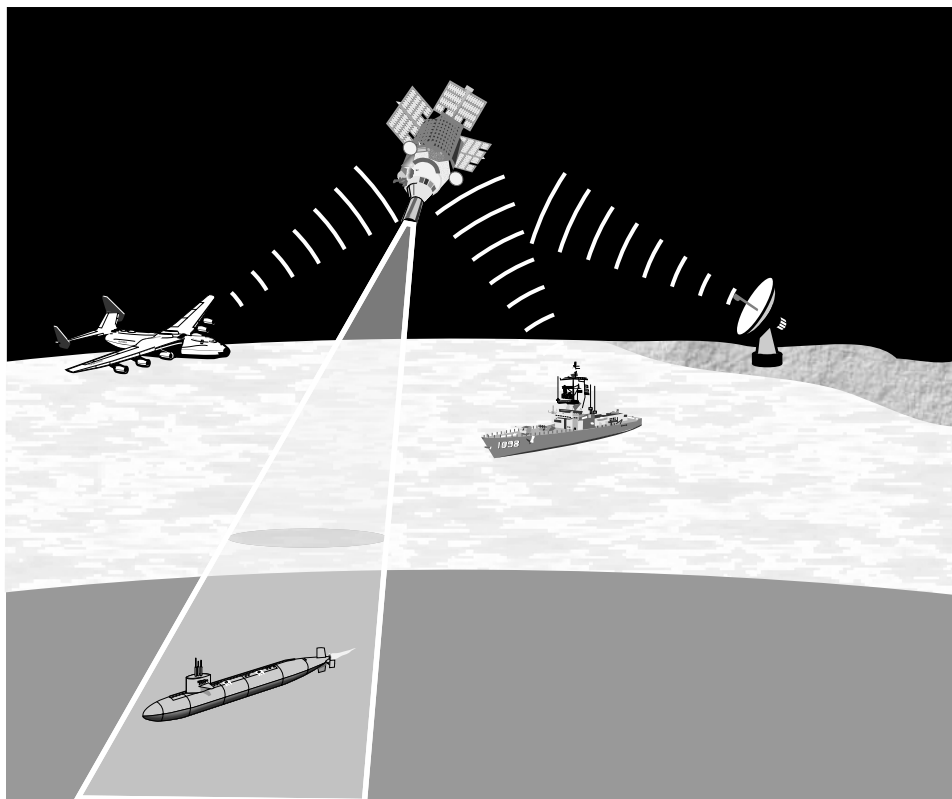
Figure 1.4: Signals are sent by conventional radio from a surface ship, ground station, or aircraft to an orbiting satellite. A blue-green laser aboard the satellite then relays the message to a submerged submarine. (Adapted with permission from Painter (1989).)

What are these challenges? Even at the Jerlov Minimum, the attenuation of seawater is not negligible, and the signal reaching a submerged vessel may be quite weak, requiring the receiver aboard the submarine to be very sensitive. This sensitivity introduces an additional complication: sunlight contains a significant blue-green component which can also penetrate the ocean and introduce noise into the received signal. One way to solve this problem is to exploit the difference between the very narrow spectral width of the blue-green laser and the much broader spectral distribution of sunlight. An optical filter with a sufficiently narrow passband can transmit most of the blue-green laser photons to the detector while rejecting most of the solar photons. In addition to a narrow passband, such a filter must have a wide field-of-view. Photons transmitted from a satellite to a submarine may pass through cloud layers that introduce scattering, and are further scattered during passage through the sea, so that they may impinge upon the submarine from a variety of
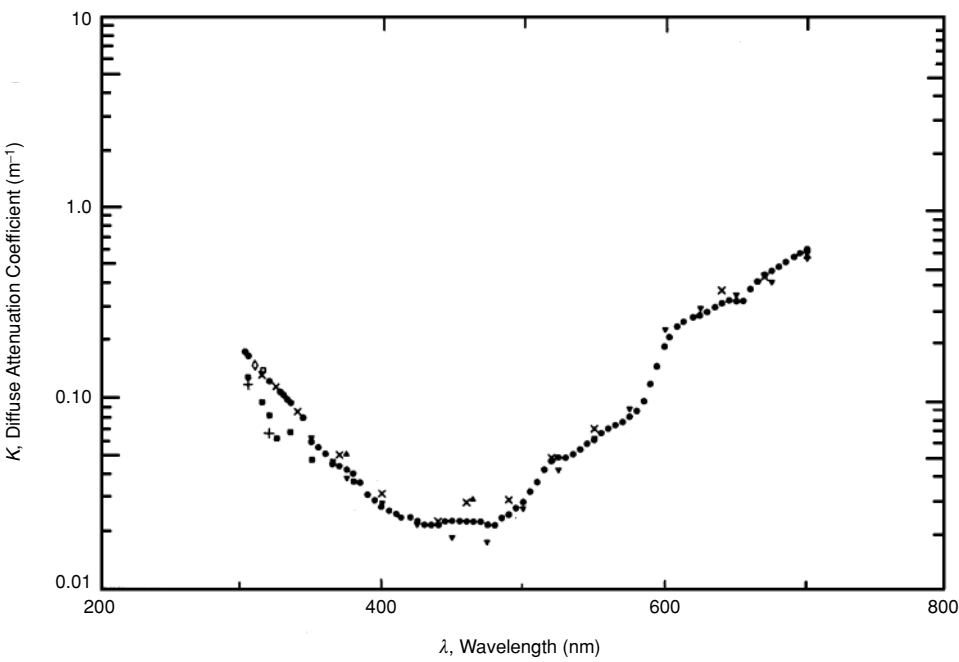
Figure 1.5: Attenuation of seawater over the blue-green portion of the optical spectrum, showing the "Jerlov Minimum" near 450 nm. Various points correspond to measurements by different authors. [Reprinted by permission from Smith and Baker (1981).]

angles. Simultanously meeting both these requirements – narrow passband and wide field-of-view – is difficult.

The most successful approach devised to meet this challenge is the "atomic resonance filter" or "ARF" (also called "QLORD"–"quantum-limited optical resonance detector"), which has the narrow passband and wide field-of-view required for submarine communications (Gelbwachs, 1988). The ratio of the spectral width $\Delta\lambda$ to center wavelength λ_0 of the passband in these filters can be $\Delta\lambda/\lambda_0 \simeq 10^{-6}$. Thus, for a center wavelength $\lambda_0 \sim 500$ nm, the width of the passband can be as narrow as ~ 0.005 Å (Marling *et al.*, 1979)! An ARF based on cesium vapor is particularly suited to submarine communications and has been pursued for this purpose. The operation of the cesium ARF can be understood from Figure 1.6. A conventional filter (e.g., colored glass such as BG-18) allows only blue-green light to enter the cesium cell. In the cesium vapor, light at 456 nm or 459 nm is absorbed to excite population from the $6s$ level to the $7p$ level. This population subsequently decays nonradiatively to the $6p$ level, through either the $7s$ or $5d$ levels. When the $6p$ population relaxes back to the ground state, infrared photons at 852 nm or 894 nm are emitted. Another conventional filter (such as RG-715 glass) permits only infrared radiation to impinge upon the detector. Since there is no overlap in the passbands of the two conventional filters, no light would reach the detector if the cesium cell