Chapter 1
Introduction to LADAR Systems

1.1 Background

RADAR (RAdio Detection And Ranging) is the process of transmitting, receiving, detecting, and processing an electromagnetic wave that reflects from a target. RADAR was first developed by the German Army in 1935.[1] As theoretical and technical developments continued, RADAR techniques and applications expanded into almost every aspect of the modern world. One area of that technical development was in the wavelength of the transmitted signal, first in the 50-cm range and later down into the millimeter and microwave regions. Pulsed light sources and optical detectors were first used in 1938 to measure the base heights of clouds.[2] The acronym LiDAR (LIght Detection And Ranging) was first used in 1953,[3] and the 1962 development of high-energy or Q-switched pulsed lasers made such sources available for LiDAR applications. In 1963, Fiocco and Sullivan published work on atmospheric observations using a ruby laser.[4] Since that time, laser-based sensors have demonstrated most, if not all, of the same functions as radio frequency (RF) or microwave RADAR. The National Institute of Standards and Technology (NIST) adopted the term LADAR (LAser Detection And Ranging) for these laser-based RADAR-type systems.[5] That term will be used in this text.

1.2 LADAR and RADAR Fundamentals

All ranging systems, whether RADAR, LiDAR, or LADAR, function by transmitting and receiving electromagnetic energy. The only difference among them is that they work in different frequency bands.[6] Therefore, many of the same considerations, such as antenna theory and propagation time, apply to all of these systems. This section will define some terms that are common to both LADAR and RADAR systems and also contrast the differences between them. Both LADAR and RADAR systems have similar configurations and subsystems, as shown in Fig. 1.1.

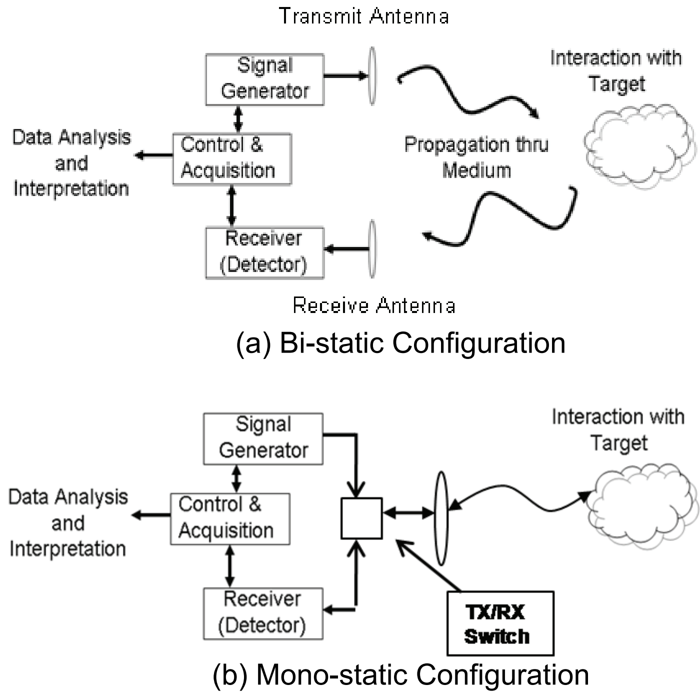

Figure 1.1 Typical LADAR/RADAR systems

The signal generator/transmitter subsystem, which is an RF/µ-wave oscillator/amplifier in RADAR and a laser (also possibly an oscillator/amplifier combination) in LADAR, determines the wavelength and waveform as well as the power of the transmitted signal. RADAR wavelengths are normally grouped in bands, and RADAR developers usually choose them in consideration of other factors such as atmospheric propagation windows. However, in principle the wavelength can be anywhere in the general band with a suitably designed oscillator and amplifier.

LADAR developers do not have that same flexibility, because they are limited to wavelengths where both suitable lasers and detectors are available. Gas lasers, such as CO_2 lasers, operate on discrete, narrow lines. And while the operating range of such a laser can be adjusted across a range of wavelengths (9 to 11 µm for CO_2), by varying the isotopes used in the gas mixture and by adjusting feedback into the cavity, the laser still operates or lases at one of the discrete laser lines. Solid state laser materials may only operate at a few (sometimes only 1) discrete lines. For example, a Nd:YAG laser will normally lase at 1.06 µm, but it can be forced, with a loss of power, to operate at 1.4 µm. While various laser materials can operate over a range of deep ultraviolet (DUV, < 250 nm) to long-wave IR (11 µm) material properties, atmospheric propagation windows and other factors limit the wavelengths of practical LADAR transmitters to a few commonly used choices (e.g., 1.06 μm with Nd:YAG, 1.5 μm for erbium-doped material, and the 9–11μm band of CO_2).

While the various waveforms used in RADAR—continuous wave (CW), amplitude or frequency modulated (AM, FM), pulsed—are also used in LADAR, the mechanism for producing these waveforms is significantly different in RADAR and LADAR. Producing the desired waveform in a RADAR transmitter could be as simple as turning the oscillator (or amplifier) on and off, or using a variable source such as a voltage-controlled oscillator into an amplifier. By comparison, the various LADAR waveforms are usually created by operating on the optical path of the laser. Q-switches that can rapidly change the output coupling of the cavity are used to dump the built-up energy stored in the cavity, producing a sharp, short pulse. Components like acousto-optical modulators are used to impress modulation on the laser output. Because the optical alignment of these components is critical, care must be taken to provide very stable bases and mounts for the optical elements. The recent development of fiber-optic-based components has made LADAR elements directly analogous to their counterparts in RADAR systems.

Once a signal is generated, it must be launched toward the target. In RADAR, this is done through an antenna. While RADAR could operate with a simple dipole-type antenna, the resulting omnidirectional beam pattern would be of minimal use, so some type of directivity is needed when using an antenna such as a multielement Yagi or dishes. The optical equivalent of the antenna in a LADAR system is the telescope (or an arrangement of optical lenses). The simplest system to implement is the bi-static configuration shown in Fig. 1.1(a). Here, separate paths and antennas are used for the transmission and receiving functions. Although this configuration is mechanically simple, it does result in a larger system package, especially for some of the longer-wavelength RADAR systems. The main advantage of this configuration is that it does not require coupling of the noise produced by the antenna's backscattering of the transmitted beam into the receiver channel. This configuration is rarely used in modern RADAR systems, but it is commonly used in LADAR systems. For most current RADAR applications, the same antenna is used for both the transmission and receiving functions (mono-static configuration, Fig. 1.1(b)). While using this configuration can reduce the size and mechanical complexity of the system, it does increase the internal circuitry and number of subsystems needed. The use of only one antenna requires the incorporation of a transmit/receive (T/R) switch. In simplest terms, this T/R switch is a three-port device with one port connected to the transmitter, one to the receiver, and one to the antenna. The switch is a direction device that routes the energy coming from one port to the next port in the rotation direction (e.g., energy coming from the transmitter is routed to the antenna, and from the antenna to the receiver) with minimal energy (noise) routed in the reverse direction. In microwave systems, this switch is called a circulator, and a magnetic field in the waveguide effects the signal rotation around the signal path. In a LADAR system, a common T/R switch uses waveplates to rotate the polarization of the laser beam and a polarizationsensitive beamsplitter to route the energy into the proper channel. Again, the fiber optic industry recently developed fiber-coupled circulators that are directly analogous to the microwave waveguide and circulator.

The final subsystem shown in Fig. 1.1 is the receiver. For both RADAR and LADAR systems, the receiver function transforms the propagating energy captured by the antenna into an electrical signal that can be processed to extract the desired information. In a RADAR system, the fluctuating electromagnetic fields of the returning signal induce currents in the receiver that can be picked up by the detector and amplified, thus creating the signal processed by subsequent subsystems. In a LADAR system, the returning photons cannot directly induce this type of current. Instead, a photodiode is used to convert the photons to current. Charge carriers are generated in response to light incident upon the photodiode, and the photon energy of received light is converted into an electrical signal by releasing and accelerating current-conducting carriers within the semiconductor. This light-induced current of the photodiode is proportional to the intensity of the incident radiation and is the signal that is transferred to other subsystems within the receiver.

Since all electromagnetic energy travels at the speed of light c, in free space, the relationship between the range R and the round-trip travel time t is given by

$$t = 2R / c. \tag{1.1}$$

Since few terrestrial LADAR applications have round-trip times that even approach seconds, time is usually accounted for in units of nanoseconds (t_{ns}). Solving Eq. (1.1) for the range yields

$$R = c \times t / 2 = t_{ns} \times 0.150\,\mathrm{m}. \tag{1.2}$$

A term that is useful for characterizing LADAR systems but is often used incorrectly is resolution. The NIST report defines resolution as "the smallest distance separation between two distinct objects illuminated by a LADAR source that can be detected in the signal return."[5] The most common misuse of the term is in reference to imaging systems to describe the size, in range, that such a system can properly image.

In Fig. 1.2, the target is a step-like structure where the range difference between the two surfaces (a, b) is ΔR. In the top figure, the beamwidth θ_t is small enough that a single surface is illuminated for each pulse (P_1, P_2). Then ΔR becomes

$$\Delta R = (R_1 - R_2). \tag{1.3}$$

In Eq. (1.3), R_1 is the distance to surface b, and R_2 is the distance to surface a. In the bottom portion of Fig. 1.2, the beamwidth is large enough that both surfaces

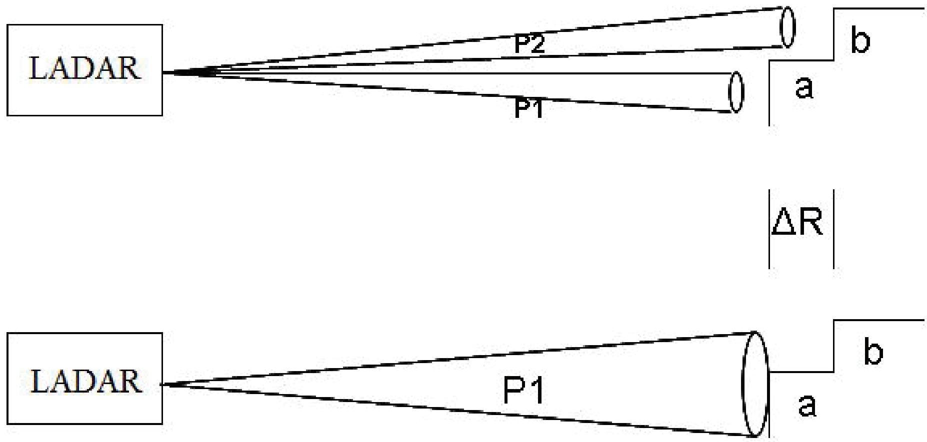

Figure 1.2 Range accuracy versus resolution.

of the target are simultaneously illuminated by a single pulse. This is the situation to which the NIST definition refers. In the former situation, we can write

$$t + \tau = 2(R + \Delta R) / c, \tag{1.4}$$

where τ is the time of flight for distance ΔR. If we subtract Eq. (1.1), we get

$$\tau = 2\Delta R / c. \tag{1.5}$$

A simple system such as a range finder usually makes only a single range measurement per transmitted pulse and does not consider range resolution. However, some systems have processors that can determine multiple ranges from a single return (i.e., “first-pulse, second-pulse” or “first-pulse, last-pulse” logic) that can be useful for applications such as removing foliage from an image to produce “bare-earth” terrain maps.

Another common aspect of LADAR and RADAR systems is antenna beamwidth, which is usually referenced as the half-power points (θ_t). In general, the beamwidth is given by

$$\theta_t = 1.22\lambda / D_t \ (in\ radians), \tag{1.6}$$

where λ is the wavelength of the light being transmitted, and D_t is the antenna diameter.[7] This assumes a uniformly illuminated aperture. Unlike range accuracy and resolution, the beamwidth, and ultimately the spatial resolution, of a system is a direct function of the signal wavelength. For example, for a 3-cm RADAR system with a 1-m antenna, θ_t = 0.03 rad (1.72 deg); and for a 1.06-µm LADAR with a 50-mm (2 in) aperture, θ_t = .000021 rad (only 0.0012 deg).

For most modern RADAR systems, the targets of interest are usually smaller than the transmitted beamwidth, and the targets act as isotropic scatterers. On the other hand, LADAR systems often have beamwidths smaller than the targets, and the targets can resemble anything from a Lambertian to a specular reflector and often are combinations of both. Another target characteristic to which RADAR and LADAR respond differently is surface roughness (more correctly, the scale of the surface roughness). Measuring this roughness can be of particular interest for applications such as monitoring sea states. Rough targets scatter the incident electromagnetic energy diffusely; smooth targets scatter specularly. Surface roughness is a relative measure and depends on the wavelength of the illuminating signal. A surface that appears rough at one wavelength might appear smooth when illuminated with longer-wavelength radiation. With RADAR, as the wavelength decreases, the size of features on the target that can be observed also decreases. At the longest wavelengths, the largest features, such as the size and location of aircraft wings, are not observable but can vary the intensity of the reflected signal as the target viewing angle changes. As the RADAR wavelength gets shorter, features such as the joints between wings and fuselage can start acting like corner cube (commonly known as retro) reflectors and can significantly contribute to the reflected signal. At the shortest wavelengths, seams between body panels—even the heads of rivets and screws—can act as retroreflectors that enhance the reflected signal. The unique shapes and construction techniques used on so-called "stealth" aircraft are attempts to counter these surface and retro effects and reduce the apparent target size to the RADAR cross section (RCS). As the wavelength decreases farther into the LADAR bands, the target surfaces themselves can produce specular and retroreflections.

Except for a few bands around 22.2, 60, and 94 GHz, most RADAR systems are not affected by the same atmospheric absorption attenuation that affects LADAR. Atmospheric attenuation is explored in more detail in Sec. 1.3.2. Since RADAR wavelengths are much longer than the diameters of most atmospheric aerosols, scattering is not an issue. One exception is at the shorter wavelengths (Ku, K bands) where weather RADAR takes advantage of scattering by rain droplets and snowflakes. In general, the following comparisons can be stated for LADAR and RADAR systems:

- Optically thick clouds and precipitation can attenuate a LADAR beam, while RADAR scatterers may consist of clouds and hydrometeors (e.g., rain or frozen precipitation). Thus, RADAR systems are generally less susceptible to atmospheric absorption effects than LADAR systems.
- LADAR beam divergence can be two to three orders of magnitude smaller than conventional 5- and 10-cm-wavelength RADAR. This gives LADAR systems superior spatial resolution but a less efficient wide-area search capability than a RADAR system.
- The combination of the short pulse (of the order of 10^{-8} sec) and the small-beam divergence (about 10^{-3} to 10^{-4} rad) creates small illuminated

volumes for LADAR (about a few m^3 at ranges of tens of km). This makes LADAR better at conducting measurements in confined places such as urban areas.

The choice of whether to use a RADAR or LADAR system for a given application is ultimately driven by mission requirements. Neither LADAR nor RADAR is necessarily a better technology, but for a specific application one may be more appropriate than the other.

1.2.1 Heterodyne versus direct detection

The differences between heterodyne (or coherent) receivers and direct-detection systems, which are incoherent in nature, are illustrated in Fig. 1.3. The primary difference between the two types of systems is that in the heterodyne receiver a portion of the outgoing laser energy is split off and redirected to the receiver detector. This energy is then aligned with that from the receiver aperture, and the detector then operates as a classical mixer. As with conventional RF receivers, the output current $i(t)$ of the detector/mixer can be written as[8]

$$i(t) = (\eta q / h\upsilon)\{P_{LO} + P_s(t) + 2[P_{LO}P_S(t)]^{\frac{1}{2}}\cos[2\pi(\upsilon_{LO} - \upsilon_S)t + \phi_S(t)]\}, \tag{1.7}$$

where υ_{LO} is the optical frequency of the local oscillator (LO) and signal, and φ_S is the relative phase between the LO and signal due to the Doppler effect. The frequency of the laser energy reflected by the target moving at a velocity (V) relative to the LADAR is shifted by

$$\Delta\upsilon = \pm(2V / c)\upsilon \tag{1.8}$$

and

$$\upsilon_S = \upsilon_{LO} + \Delta\upsilon. \tag{1.9}$$

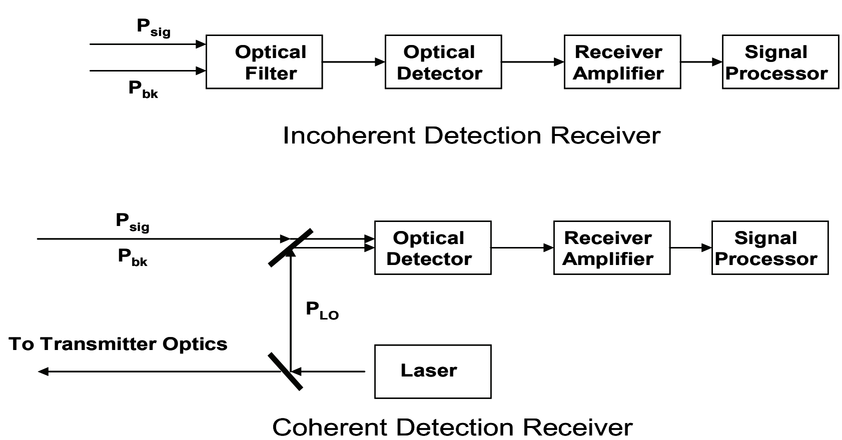

Figure 1.3 Coherent versus direct-detection receivers.

The last term in Eq. (1.7) is the signal of interest in a heterodyne system, with the envelope of the signal following the shape of $P_s(t)$ and the signal within the envelope varying at $\Delta\upsilon$ [now referred to as the intermediate frequency (IF)]. While the frequency of the optical fields is too high for any electronic circuits to respond to, modern electronics can easily accommodate the IF in a heterodyne system [for example, a 1.06-μm system would have an IF of approximately 1 MHz. While a heterodyne system has a theoretical advantage over direct detection of about 30 dB,[9] this is rarely achieved in a practical system. Effects such as phase front distortion due to the target's propagation and depolarization of the signal can significantly reduce the heterodyne efficiency of the receiver.[10] A coherent detection system is often more complex than a direct-detection LADAR and requires the use of diffraction-limited optics to achieve maximum efficiency.

1.3 LADAR Range Equation

Various types of LADAR systems take advantage of different portions of the signal propagation process. The description of this process, the LADAR equation, is directly analogous to the original RADAR equation and can be broken down into several terms that quantify the contribution or effects of various elements of the process illustrated in Fig. 1.1. The range equation is widely used as an analytical tool for computing the power received (P_r) from a target illuminated by a laser pulse containing a given power (P_t). The range equation has many forms that depend on the physical process being interrogated by the LADAR. In general, the LADAR equation covers the following aspects of laser propagation, reflection, and reception:

- Laser transmitted power P_t
- Laser transmitter beam diameter and angular divergence θ_t
- Atmospheric transmission τ_a
- Target surface reflectivity ρ_t
- Target surface angular dispersion θ_R
- LADAR receiver quantum efficiency η

1.3.1 Laser transmitter models

For purposes of this chapter, we will assume the laser transmitter used in a typical direct-detection LADAR system fires a pulse of laser energy that exists for a short period of time. In later chapters, the shape of the pulse in time or the waveform will be considered, but for this elementary analysis, we shall consider only a pulse that is rectangular in shape and exists for a period of time equal to the pulse width (typically in nanoseconds, ns). The instantaneous power in watts (W) transmitted by the system P_t is then the energy in the pulse divided by the pulse width in time. This describes the temporal shape of the power output of the transmitter.