

Field Guide to **Lidar**

Paul McManamon

SPIE Field Guides
Volume FG36

John E. Greivenkamp, Series Editor

SPIE PRESS
Bellingham, Washington USA

Library of Congress Preassigned Control Number:
2015931195

for

Field Guide to Lidar (ISBN 9781628416541).

Published by

SPIE

P.O. Box 10

Bellingham, Washington 98227-0010 USA

Phone: +1.360.676.3290

Fax: +1.360.647.1445

Email: books@spie.org

Web: <http://spie.org>

Copyright © 2015 Society of Photo-Optical Instrumentation
Engineers (SPIE)

All rights reserved. No part of this publication may be
reproduced or distributed in any form or by any means
without written permission of the publisher.

The content of this book reflects the work and thought of
the author. Every effort has been made to publish reliable
and accurate information herein, but the publisher is not
responsible for the validity of the information or for any
outcomes resulting from reliance thereon.

Printed in the United States of America.

First printing

SPIE.

Introduction to the Series

Welcome to the *SPIE Field Guides*—a series of publications written directly for the practicing engineer or scientist. Many textbooks and professional reference books cover optical principles and techniques in depth. The aim of the *SPIE Field Guides* is to distill this information, providing readers with a handy desk or briefcase reference that provides basic, essential information about optical principles, techniques, or phenomena, including definitions and descriptions, key equations, illustrations, application examples, design considerations, and additional resources. A significant effort will be made to provide a consistent notation and style between volumes in the series.

Each *SPIE Field Guide* addresses a major field of optical science and technology. The concept of these *Field Guides* is a format-intensive presentation based on figures and equations supplemented by concise explanations. In most cases, this modular approach places a single topic on a page, and provides full coverage of that topic on that page. Highlights, insights, and rules of thumb are displayed in sidebars to the main text. The appendices at the end of each *Field Guide* provide additional information such as related material outside the main scope of the volume, key mathematical relationships, and alternative methods. While complete in their coverage, the concise presentation may not be appropriate for those new to the field.

The *SPIE Field Guides* are intended to be living documents. The modular page-based presentation format allows them to be updated and expanded. We are interested in your suggestions for new *Field Guide* topics as well as what material should be added to an individual volume to make these *Field Guides* more useful to you. Please contact us at fieldguides@SPIE.org.

John E. Greivenkamp, *Series Editor*
College of Optical Sciences
The University of Arizona

The Field Guide Series

Keep information at your fingertips with the SPIE Field Guides:

Adaptive Optics, Second Edition, Robert Tyson & Benjamin Frazier
Atmospheric Optics, Larry Andrews
Binoculars and Scopes, Paul Yoder, Jr. & Daniel Vukobratovich
Diffraction Optics, Yakov Soskind
Digital Micro-Optics, Bernard Kress
Displacement Measuring Interferometry, Jonathan D. Ellis
Fiber Optic Sensors, William Spillman, Jr. & Eric Udd
Geometrical Optics, John Greivenkamp
Holography, Pierre-Alexandre Blanche
Illumination, Angelo Arecchi, Tahar Messadi, & John Koshel
Image Processing, Khan M. Iftakharuddin & Abdul Awwal
Infrared Systems, Detectors, and FPAs, 2nd Edition, Arnold Daniels
Interferometric Optical Testing, Eric Goodwin & Jim Wyant
Laser Pulse Generation, Rüdiger Paschotta
Lasers, Rüdiger Paschotta
Lens Design, Julie Bentley & Craig Olson
Linear Systems in Optics, J. Scott Tyo & Andrey Alenin
Microscopy, Tomasz Tkaczyk
Nonlinear Optics, Peter Powers
Optical Fabrication, Ray Williamson
Optical Fiber Technology, Rüdiger Paschotta
Optical Lithography, Chris Mack
Optical Thin Films, Ronald Willey
Optomechanical Design and Analysis, Katie Schwertz & James Burge
Physical Optics, Daniel Smith
Polarization, Edward Collett
Probability, Random Processes, and Random Data Analysis, Larry C. Andrews & Ronald L. Phillips
Radiometry, Barbara Grant
Special Functions for Engineers, Larry Andrews
Spectroscopy, David Ball
Terahertz Sources, Detectors, and Optics, Créidhe O'Sullivan & J. Anthony Murphy
Visual and Ophthalmic Optics, Jim Schwiegerling

Field Guide to Lidar

This *Field Guide* covers active electro-optical sensing, in which a sensor sends out a laser pulse and then measures the parameters of the return signal. Various groups refer to this type of sensor as a ladar, lidar, LIDAR, LADAR, or laser radar. For simplicity, only the term lidar is used throughout this book.

The book is presented from the perspective of a lidar engineer. It covers a wide breadth, from simple 2D direct-detection lidars to multiple subaperture synthetic aperture lidars. It also covers a broad range of objects to be viewed, and distances from which to view the objects. Lasers and modulation are discussed in the context of their use in lidars. Other topics covered include receivers, apertures, and atmospheric effects in the context of lidar use and design.

All lidars will be limited by the media between the lidar and the target, but atmospheric compensation techniques can often mitigate this limitation. These limitations and compensation approaches are presented. Many types of lidars are included along with appropriate data processing techniques. The lidar range equation in its many variations is discussed along with receiver noise issues that determine how much signal must be received to detect an object.

This *Field Guide* is a handy reference to quickly access information on any aspect of lidars. It will be useful to students and lidar scientists or engineers who need an occasional reminder of the correct approaches or equations to use in certain applications. It will also be useful to systems engineers gaining a perspective on this rapidly growing technology.

Paul McManamon
March 2015

Table of Contents

Glossary of Symbols and Acronyms	x
Introduction	1
Introduction	1
Terms for Active Electro-optic Sensing	2
Types of Lidars	3
Lidars for Surface-Scattering (Hard) Targets	4
Lidars for Volume-Scattering (Soft) Targets	5
History of Lidar	6
Lidar Detection Modes	7
Spatial Coherence	8
Temporal Coherence	9
Eye Safety Considerations	10
Laser Safety Categories	11
Monostatic versus Bistatic Lidar	12
Transmit/Receive Isolation	13
Lidar Range Equation	14
Lidar Range Equation	14
Lidar Cross Section	15
Cross Section of a Corner Cube	16
Speckle	17
Atmospheric Absorption	18
Atmospheric Scattering	19
Atmospheric Turbulence	20
Aero-optical Effects on Lidar	21
Extended (Deep) Turbulence	22
Laser Power for Lidar	23
Lidar Signal-to-Noise Ratio	24
Direct Detection Signal-to-Noise Ratio	25
Noise Probability Density Functions	26
Thermal Noise	27
Shot Noise	28
The Sun as Background Noise	29
Dark Current, $1/f$, and Excess Noise	30
Avalanche Photodiodes and Direct Detection	31
Number of Photons Required for a GMAPD Lidar	
Camera	32
Heterodyne Detection	33

Temporal Heterodyne Detection	34
Heterodyne Mixing Efficiency	35
Quadrature Detection	36
Carrier-to-Noise Ratio for Temporal Heterodyne Detection	37
Spatial Heterodyne Detection/Digital Holography	38
SNR for Spatial Heterodyne Detection	39
Types of Lidars	40
1D Range-Only Lidar	40
Tomographic Imaging Lidar	41
Range-Gated Active Imaging (2D Lidar)	42
3D Scanning Lidar	43
3D Flash Imaging	44
Geiger-Mode APD Flash Lidar	45
Linear-Mode APD Flash Lidar	46
Polarization-based Flash Lidar using Framing Cameras	47
Laser Vibration Detection	48
Synthetic Aperture Lidar	49
Inverse Synthetic Aperture Lidar	50
Range Doppler Imaging Lidar	51
Laser-Induced Breakdown Spectroscopy	52
Laser-Induced Fluorescence Lidar	53
Active Multispectral Lidar	54
Lidars Using Polarization as a Discriminant	55
Speckle Imaging Lidar	56
Phased Array of Phased-Array Imaging Lidar	57
Multiple Subapertures on Receive for Lidar	58
Multiple-Input, Multiple-Output Lidar	59
Methods of Phasing MIMO Lidars	60
Lidar Sources and Modulations	61
Lidar Sources and Modulations	61
Laser Resonators	62
Three-Level and Four-Level Lasers	63
Bulk Solid State Lasers for Lidar	64
Fiber Lasers for Lidar	65
Higher-Peak-Power Waveguide Lasers for Lidar	66
Diode Lasers for Lidar	67

Quantum Cascade Lasers for Lidar	68
Laser Pumping Considerations	69
Nonlinear Devices to Change the Lidar Wavelength	70
Q-Switched Lasers for Lidar	71
Pockels Cells	72
Mode-Locked Lasers for Lidar	73
Laser Seeding for Lidar	74
Laser Amplifiers for Lidar	75
Multiple Coherent Laser Transmitters	76
Laser Waveforms for Lidar	77
Polypulse Laser Waveforms	78
Linear Frequency Modulation for Lidar	79
Pseudo-random-Coded Lidar	80
RF Modulation of a Direct Detection Lidar	81
Lidar Receivers	82
Linear-Mode APD Arrays for Lidar	82
Geiger-Mode APD Arrays for Lidar	83
Receivers for Coherent Lidars	84
Acousto-optic Frequency Shifting	85
Long-Frame-Time Framing Detectors for Lidar	86
Gated Framing Cameras for 2D Lidar Imaging	87
Lidar Image Stabilization	88
Range Resolution of Lidar	89
Velocity Resolution of Lidar	90
Unambiguous Range	91
Point Spread Function	92
Beam Steering for Lidars	93
Gimbals for Use with Lidar	93
Fast-Steering Mirrors	94
Risley Prisms and Gratings	95
Rotating Polygonal Mirrors	96
Modulo 2π Beam Steering	97
Largest Steering Angle for an Optical Phased Array	98
Liquid Crystal Optical Phased Arrays	99
LC Fringing-Field Effect on Steering Efficiency	100
Reduction in Steering Efficiency Due to	
Quantization	101
Chip-Scale Optical Phased Arrays	102
MEMS Beam Steering for Lidar	103

Electrowetting Beam Steering for Lidar	104
Steerable Electro-evanescent Optical Refractors	105
Electro-optical Effects	106
Polarization Birefringent Grating Beam Steering	107
Step Angle Steering with LC Polarization Gratings	108
Multiple-Stage LCPGs	109
Lenslet-based Beam Steering	110
Electronically Written Lenslets	111
Mixed-Lenslet Arrays	112
Holographic Gratings for Beam Steering	113
Geometrical Optics	114
Lidar Processing	115
Inertial Measurement Units	115
Microscanning of Lidar Images for Improved Sampling	116
Range Measurement Processing	117
Nyquist Sampling a Range Profile	118
Threshold, Leading Edge, and Peak Detectors	119
Range Resolution, Precision, and Accuracy	120
Fourier Transforms	121
Developing 3D Maps from Lidar	122
3D Metrics for Lidar Images	123
Multiple-Subaperture Spatial Heterodyne Processing	124
Definitions of Lidar Data Processing Stages	125
Processing Laser Vibrometry Data	126
Target Classification Using Lidar	127
Equation Summary	128
Figure Sources	138
Bibliography	141
Index	143

Glossary of Symbols and Acronyms

α	amplitude of the (super) Gaussian
A	length of one side of a tetrahedral
A_{illum}	area illuminated by the transmitter
AO	acousto-optic
AOM	acousto-optic modulator
A_p	area of the pixel at the target location
APD	avalanche photodiode
APS	active-pixel sensor
A_{rec}	area of the receiver aperture
b	zero position, or offset, of the (super) Gaussian beam
B	bandwidth
c	Gaussian, or super-Gaussian, beam width
c	speed of light
cw	continuous wave
C_1	coherence length
CCD	charge-coupled device
CDMA	code-division multiple access
CMOS	complementary metal-oxide semiconductor
CNR	carrier-to-noise ratio
d	cross-range resolution
d	required lens thickness
d	width of the individual radiator or receiver
D	aperture diameter
D_{Airy}	diameter out to the zeros of the diffraction- limited spot at the focus for a circular aperture
DAS	detector angular subtense
DFLC	dual-frequency liquid crystal
DIAL	differential absorption lidar
DM	deformable mirror
DOP	degree of polarization
e	charge on an electron
E	energy at range
E_0	initial energy before traveling through the atmosphere
EBAPS [®]	electron-bombarded active-pixel sensor
EBS	electron-bombarded semiconductor
E_{in}	input electric field into a Jones matrix

Glossary of Symbols and Acronyms

E_{LO}	local oscillator field
EM	electromagnetic
EO	electro-optic
E_{out}	input electric field into a Jones matrix
E_{p}	energy in a photon
E_{R}	received energy per pulse
E_{sig}	returned signal field
E_{T}	transmitted energy per pulse
E_{th}	thermal energy
Ex_{in}	x portion of the input electric field
Ex_{out}	x portion of the output electric field
Ey_{in}	y portion of the input electric field
Ey_{out}	y portion of the output electric field
f	focal length of the lens
$f/\#$	F-number of an optical element
f_{l}	focal length of a lenslet
$f(x)$	Gaussian or super-Gaussian beam profile in one dimension
F	excess noise factor associated with the preamplifier gain
FDMA	frequency-division multiple access
FFT	fast Fourier transform
FLC	ferroelectric liquid crystal
FLIR	forward-looking infrared (camera)
FM	frequency modulated
FOV	field of view
FPA	focal plane array
FSM	fast-steering mirror
G	avalanche gain
GIQE	general image quality equation
GMAPD	Geiger-mode avalanche photodiode
GML	Geiger-mode lidar
h	Planck's constant
HWP	half-wave plate
i_{bk}	background current
i_{dk}	dark current
i_{n}	noise current in the detector
i_{s}	signal current in the detector
i_{shotLO}	shot noise from the local oscillator

Glossary of Symbols and Acronyms

$i_{\text{shot,sig}}$	shot noise from the signal
i_{th}	thermal noise current
I	intensity of the beat between the local oscillator and the return signal
I_{dkb}	bulk dark current
I_{dks}	surface dark current
IF	intermediate frequency
IMU	inertial measurement unit
IR	infrared
k	effective elastic constant
k	number of photons in M events
k	Boltzmann constant
L	distance flown
L	length of the laser cavity
LCPG	liquid crystal polarization grating
LFM	linear frequency modulation
LIBS	laser-induced breakdown spectroscopy
LIF	laser-induced fluorescence
LIMAR	laser imaging and ranging
LMAPD	linear-mode avalanche photodiode
LO	laser oscillator
LWIR	long-wave infrared
L_{λ}	radiance per wavelength
M	number of events
M^2	measure of the spatial coherence of a laser beam. An M^2 of 1 means it is diffraction limited.
MEMS	micro-electro-mechanical system
MIMO	multiple input, multiple output
MO	master oscillator
MPE	maximum permissible exposure
MWIR	midwave infrared
n	index of refraction
n	number of individual radiators or receivers
n_m	diffraction efficiency of the m^{th} order
N	number of photons per pixel received during a measurement time
N	super-Gaussian beam number. Higher numbers mean a more flat-topped beam shape.

Glossary of Symbols and Acronyms

NA	numerical aperture
NEPh	noise-equivalent photons
NIIRS	National Imagery Interpretability Rating Scale
NIR	near infrared
OPA	optical parametric amplifier
OPA	optical phased array
OPD	optical path difference
OPO	optical parametric oscillator
$p(k)$	Gaussian probability distribution
P	number of modes
PAPA	phased array of phased arrays
P_{LO}	local oscillator power
PPLN	periodically poled lithium niobate
P_S	signal power received
PSD	power spectral density
PSF	point spread function
P_T	power transmitted
P_{thdbm}	thermal noise power
q	Poisson distribution probability
q	number of discrete steps
QCL	quantum cascade laser
QWP	quarter-wave plate
r_0	Fried parameter
R	range to the target
$\Re$	detector responsivity
RF	radio frequency
R_L	load resistance
ROIC	readout integrated circuit
$R_{unambig}$	unambiguous range
$S'_3 = S_3/S_0$	normalized Stokes parameter corresponding to ellipticity of incident light
SNR	signal-to-noise ratio
SPGD	stochastic parallel gradient descent
SS	solid state
SWIR	short-wave infrared
t	cell thickness
$t_{lens}(w_{az}, w_{el})$	lens phase profile
T	temperature

Glossary of Symbols and Acronyms

T	time separation between pulses
TDMA	time-division multiple access
T_m	time period over which a measurement is made
v	velocity of the lidar with respect to the surrounding air
V	platform velocity
V	relative velocity between the lidar and the target
V	voltage on an electrode
VCSEL	vertical-cavity surface-emitting laser
V_t	threshold voltage
$W_{az}^2 + W_{el}^2$	beam width in azimuth and elevation for a Gaussian profile
β	angle between the slow axis of the half-wave plate and the x axis in the Jones matrix
β	atmospheric decay constant
γ	viscosity
Δf	change in frequency due to the Doppler shift
Δn	change in index of refraction
Δz	surface roughness
ΔR	range resolution
Δt	mode-locked pulse width
ΔV	velocity resolution
Δx	lenslet motion
$\Delta \theta$	angular resolution for a synthetic aperture lidar
$\Delta \lambda$	linewidth of the laser in wavelength
$\Delta \phi$	angular motion used in an inverse synthetic aperture lidar image
η	steering efficiency due to quantization error
η_{atm}	transmission of the atmosphere in one direction
η_h	heterodyne mixing efficiency
η_{sys}	total transmission of the lidar system, both in and out
θ	angular motion created by the lenslet
θ_{max}	maximum steering angle

Glossary of Symbols and Acronyms

ϑ	angle of deflection for an AO modulator
ϑ	full beam width, half maximum diffraction limit
λ	wavelength
λ_i	wavelength of the idler laser
λ_p	wavelength of the pump laser
λ_s	wavelength of the signal laser
Λ	acousto-optical wavelength
Λ	width between resets
Λ_F	width of the flyback region
ν	carrier frequency of light ($\omega = 2\pi\nu$)
ρ	radius of the microlens
ρ_t	reflectance of the area
σ	cross section
τ_0	coherence time
τ_d	time required to return to no-voltage state
τ_m	mode-locked pulse separation
φ	phase retardation of the half-wave plate
ω_{sig}	frequency (in radians) of the return signal
ω_{LO}	frequency (in radians) of the local oscillator

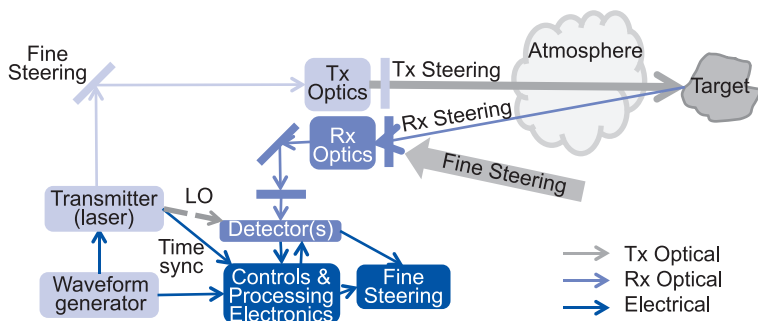
Introduction

Lidar is referred to by many names, but all of them mean obtaining information by bouncing light off of an object. A **laser signal** is sent out, and information is gathered based on the reflection from objects.

A **waveform generator** generates a **laser waveform**. The setup can include a single laser or a master oscillator with multiple lasers or laser amplifiers. Transmit optics are also needed. The laser is aimed at the target of interest. Light must traverse a media, usually atmosphere, to a target. It bounces off the target and traverses the media again until it is captured by receive optics. The receive optics must also view the target. The returned light hits a detector or a detector array. With a detector array, an image can be made with a single pulse. This is called **flash imaging**.

The setup can also include a local oscillator that beats against the return signal on the detector to allow measurement of phase as well as intensity. This is called **coherent lidar**. A timing signal from the transmitter is needed to determine range by measuring time of flight of the laser pulse to and from the target. The signal generated by the detector is digitized and then processed to make an image, or to generate information such as velocity or vibration based on the return Doppler shift.

In **monostatic lidar**, light is emitted and received by the same aperture. The figure shows two separate apertures, one for transmit and one for receive, in a setup referred to as **bistatic lidar**.

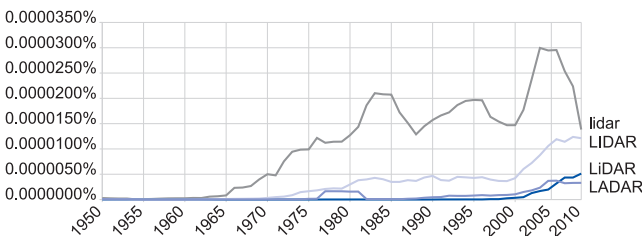


Terms for Active Electro-optic Sensing

This book uses the term *lidar* because lidar is currently the most commonly used term, but other terms for active electro-optic (EO) sensing are also in use:

- **LIDAR**: LIght Detection And Ranging
 - Used by the National Geospatial Agency for active EO 3D imaging and mapping.
 - Most commonly used for commercial applications.
 - Historically used with volume scattering targets such as atmosphere, or for chemical vapor detection.
- **LADAR**: LAser Detection And Ranging
 - Historically used with surface-scattering targets.
 - Adopted by NIST as the standard term for active EO sensing.
- **Laser radar** and **laser remote sensing** are also sometimes used.
- For reference, radar stands for RAdio Detection And Ranging.

This book uses the term *lidar* as both a technology and a specific instrument. Historically, lidar has been used by the active EO sensing community in conjunction with measuring volume-based targets (soft targets), such as aerosols, and ladar has been used in conjunction with surface-based reflections (hard targets). Lidar has been used more often in commercial applications, and therefore has the more widespread use. Capitalization varies; the most popular version is lidar, followed by LIDAR, then ladar. The various terms are often used almost interchangeably, depending on who is using them. This can cause confusion. The chart shows usage of the various terms since 1965.

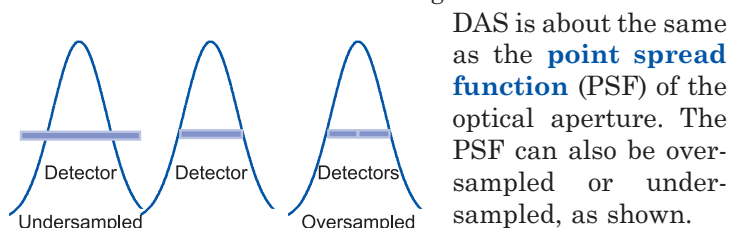


Types of Lidars

Range in a lidar is measured by timing how long it takes light to hit an object and return. **Range resolution** ΔR is given by

$$\Delta R = \frac{c}{2B}$$

where c is the speed of light in the intervening medium, and B is the system bandwidth, typically given by the transmit signal bandwidth. Angle information is detected by one detector or an array of detectors sampling the **field of view** (FOV). Angular resolutions can be limited by either the **detector angular subtense** (DAS) or by diffraction. The center of the figure shows where the



Range-only lidar, or **1D lidar**, measures only one dimension: range. **2D lidar** is similar to passive imaging in that it measures both azimuth and elevation but does not measure range. A common type of lidar is the **3D lidar**, measuring angle/angle/range. A 3D image can be only 3D, with no grayscale, or can measure grayscale. Polarization and color (wavelength) are other dimensions that can be measured.

A **coherent lidar**, which measures phase and amplitude of the return signal, can directly measure the velocity of an object by measuring Doppler shift in frequency. Angle measurement can be made using an aperture that is synthesized by sampling the pupil plane field in many locations and then constructing a larger pupil plane field sample, which can be Fourier transformed to obtain the image.

Synthetic aperture lidar (SAL) is one form of a synthesized lidar aperture. Other forms can include multiple physical apertures on receive only, or on both receive and transmit.

Lidars for Surface-Scattering (Hard) Targets

Lidar can be used for many different applications because it is such a robust sensing approach. Here is a small sample of its vast and growing applications.

3D lidar is used for **3D mapping**. 2D lidar will look a lot like conventional EO images. Velocity and vibration can be measured with 1D lidar. A well-known application of lidar is the driverless car. All of the DARPA Grand Challenge finishers used at least one lidar. Google uses a lidar on their driverless car. Police speed detection is a common lidar use. Lidar mounted in a handheld camera can pick out one vehicle in a cluttered traffic scene. The Microsoft Kinect game system is a new very widespread use of a simple lidar.

Forestry lidar is unique in its ability to measure the vertical structure of forest canopies, map the ground beneath the forest, and estimate canopy bulk density. Lidar can map flood plains and the ground in coastal areas in and out of the water. Transportation corridors can be 3D mapped to support engineering high-scale mapping accuracy. The military uses 3D map data for **route planning** and for **precise object identification**.

Lidar's high accuracy means that a quick survey will give precise volumetric measurements for **oil and gas exploration** and for quarrying. Lidar can provide the data for **cellular network planning** to determine line of sight and view-shed for a proposed cellular antenna. It allows any physical object to be re-created in a computer environment or with a 3D printer. Lidars make recording the scene of accidents and crimes quick, easy, and precise. Lidar has been used in **archeological mapping** and is a useful tool when designing and constructing new buildings.

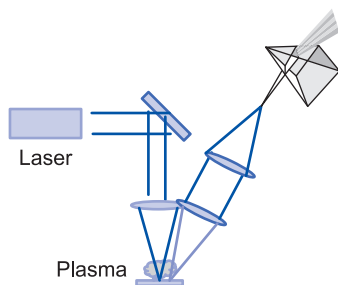
Lidar enables **surveys** to be taken of places that might be considered too dangerous for humans to enter. A robotic vehicle can be sent down into sewers to take detailed surveys of the interior of the system.

Lidars for Volume-Scattering (Soft) Targets

Wind speed can be measured either by the Doppler shift along the path from the lidar or from multiple accurate range measurements. **Doppler lidar** measures wind speed along the beam by measuring the frequency shift of the backscattered light for military applications such as air drop of cargo, or to map the wind fields around an airport or a set of wind turbines. **Differential absorption lidar** (DIAL) is used to detect range-resolved trace amounts of gases such as ozone, carbon dioxide, or water vapor. The lidar transmits two wavelengths: an online wavelength that is absorbed by the gas of interest and an offline wavelength that is not absorbed. The differential absorption between the two wavelengths is a measure of the concentration of the gas.

Raman lidar is also used for measuring the concentration of atmospheric gases but exploits inelastic scattering to single out the gas of interest. A small portion of the energy of the transmitted light is deposited in the gas during the scattering process, shifting the scattered light to a longer wavelength by an amount that is unique to the species of interest.

Laser-induced fluorescence (LIF) lidars are similar to Raman lidars in that they only require that the entity being detected be present in the atmosphere or on the surface of a target of interest. Unlike Raman lidars, the LIF systems seek to create real transitions in the entity through excitation to higher electronic levels.



sensor to remotely determine rock type. The figure shows a schematic of a LIBS lidar.

In this book, **laser-induced breakdown spectroscopy** (LIBS) appears on the volume-scattering page because first a small amount of material is vaporized, then spectroscopy is performed on the vaporized material. The Curiosity spacecraft on Mars uses a LIBS

History of Lidar

Lidar has a rich history of about 50 years. While some early work was done before the invention of the laser, most **lidar development** started shortly after this invention. In the 1960s, a ruby laser was bounced off of the moon to accurately measure range.

Initially, lidars used various lasers, until **CO₂** became the popular choice for coherent lidar, and **Nd:YAG** for laser range finders and designators. CO₂ reigned as the laser of choice for coherent lidar from the early 1970s until the 1980s or 1990s. Most CO₂ lidars were at 10.6- μ m wavelength, although, to avoid atmospheric CO₂ absorption, some CO₂ lidars used different isotopes. CO₂-based navigation lidar was deployed in the 1980s and 1990s, but with the advent of GPS became less popular.

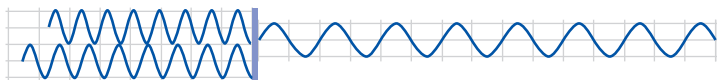
The exceptions to laser dominance in early lidar development being CO₂ were laser designators, which can be considered as a form of bistatic lidar, and **laser range finders**. Nd:YAG, at 1.064- μ m wavelength, was the laser of choice for both.

Laser designators started to be developed in the late 1960s and are still being used today. Initially, Nd:YAG lasers were flashlamp pumped. In the 1980s to 1990s, multiple solid state lasers came into their own for lidar application, including coherent solid state lidars. **Laser diodes** also became prominent, allowing a reliable and efficient method of pumping. Wind sensing, navigation, terrain following, 2D and 3D imaging, and velocity detection are some of the other lidar uses that have been pursued.

Lidar Detection Modes

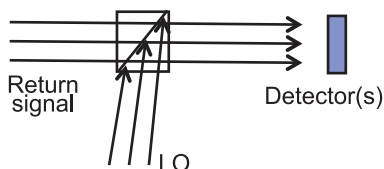
An optical detector only responds to the intensity of light. It cannot follow the rapid fluctuations of the 200- to 600-THz carrier frequencies of 0.5- to 1.5- μm light. When light hits a detector, it generates a voltage equal to the square of the intensity I of the impinging light. In ideal **direct detection**, only the lidar return hits the detector, causing a response equal to the square of the impinging light intensity. In **coherent lidar**, the return signal beats against a sample of the emitted signal, called the **local oscillator** (LO). In this case, the detector can respond to the beat (or difference) frequency between the return signal and the LO field:

$$I = 2E_{\text{sig}} \cdot E_{\text{LO}} \cdot \exp[-jt(\omega_{\text{sig}} - \omega_{\text{LO}})]$$



In **temporal heterodyne detection**, the LO and return signals are spatially aligned to interfere on the detector:

If they are not aligned, high-spatial-frequency fringes will lower heterodyne mixing efficiency. The frequency of the LO is offset so that the direction of target velocity can be determined, and to reduce **1/f noise**. If there is no frequency offset, temporal heterodyne detection is called **homodyne detection**.



In **spatial heterodyne detection**, also called **digital holography**, the LO and the return signal are slightly misaligned, as shown in the figure. The tilt

between the LO and the return signal creates fringes that can be sampled by a detector array. If the angular tilt is too large, the fringes will be too rapid and will average out across a detector. In **phase shift interferometry**, the reflected wave is mixed with a like-frequency LO. The interferogram is measured on a 2D detector array. Multiple interferograms are recorded in which a known piston phase has been applied to the LO. Three or more interferograms are processed to extract the field amplitude and phase.

Spatial Coherence

Spatial coherence relates to the width of the laser beam. For lidars that have a single detector, the laser beam needs to illuminate only the area covered by the single detector. If the lidar uses a detector array (flash lidar), the laser illuminates an area as large as the detector array, so the laser beam can be wider. The **diffraction limit** provides the smallest possible **laser beam divergence**. The full beam width, half maximum diffraction limit is

$$\vartheta \approx 1.03\lambda/D$$

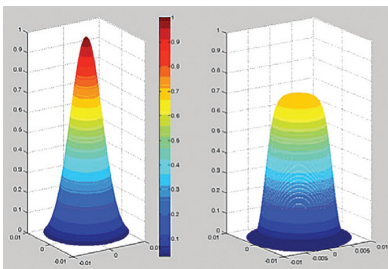
where D is the diameter of the transmit aperture. If the transmit aperture is the same size as the receive aperture, high spatial coherence is not needed for flash imaging, which is shown in the figure.



Many lidar illumination beams are **Gaussian**, but a flat-topped beam is often desirable. A **super-Gaussian** beam has a flatter top than a Gaussian beam, allowing a more uniform illumination pattern. The equation for a Gaussian beam if $N = 2$, and for a super-Gaussian beam if N is larger than 2, is

$$f(x) = ae^{-\frac{(x-b)^N}{2c^2}}$$

where a is the magnitude of the Gaussian beam, b is the offset from zero, and c is the width of the Gaussian beam. The figure below shows a Gaussian beam and a super-Gaussian shape. With a finite-sized transmit aperture, less energy can be clipped from



the laser beam with any flatter-topped beam shape, including various-powered super-Gaussian shapes. A Gaussian beam concentrates more of its energy in the middle of the **point spread function (PSF)**, compared to a super-Gaussian, or flat top, beam.

Temporal Coherence

Temporal coherence indicates the degree of **monochromaticity** of a laser beam. A lidar that beats the return signal against a LO (a heterodyne, or coherent, lidar) usually uses a narrow linewidth for the beating between lasers to produce interference.

Coherence length of a laser is a measure of the distance the laser beam travels before a full wavelength of phase shift occurs between the extreme frequencies in the laser linewidth. It is impossible to obtain interference, or heterodyne mixing, beyond the coherence length. Without delaying a copy of the emitted signal, the coherence length of a coherent lidar must be greater than the round-trip distance to the target. If a sample of the outgoing signal is delayed, the coherence length in a coherent lidar must be more than twice the range depth of the target. The coherence length is the speed of light divided by linewidth:

Laser Bandwidth	Laser Coherence Length
1 kHz	300 km
10 kHz	30 km
100 kHz	3 km
1 MHz	300 m
10 MHz	30 m
100 MHz	3 m
1 GHz	30 cm
10 GHz	3 cm

$$C_1 = c/B$$

If two laser beams are beat against each other with less than one coherence length separation in the distances travelled, a beat will occur between the two beams. With more than one coherence length of travel before interfering, the beams create **random interference**, and phase information is lost. **Speckle** is a form of interference, so narrow-linewidth lasers exhibit speckle. It is useful to compare laser linewidth in wavelength to laser linewidth in frequency:

$$B = \frac{c}{\lambda^2} \Delta\lambda$$

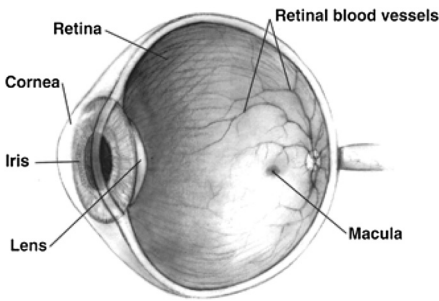
$\Delta\lambda$ (nm)	Δf (GHz)
0.1	13
1	126
10	125
100	12650

In general, laser filters are specified in wavelength, but for heterodyne lidars the bandwidth is quoted in frequency. A very narrow filter on receive does not usually result in as narrow a linewidth as in heterodyne detection.

Eye Safety Considerations

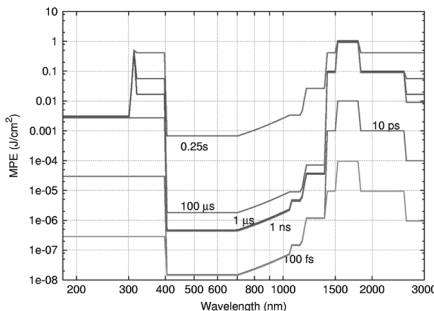
Laser radiation can damage the eye by burning the **retina** after magnification or by burning the **surface of the eye**. The eye has significant magnification when it focuses light on the retina, so the **damage threshold** for light entering the eye and focusing on the retina is much lower than the damage threshold for burning the surface of the eye. Lasers beyond $\sim 1.5\ \mu\text{m}$, or below about $0.4\ \mu\text{m}$, are safer than lasers between these wavelengths because water in the eye

absorbs wavelengths in these regions, preventing light from focusing on the retina. These lasers can still cause burns on the surface of the eye but do not create the large increase in irradiance resulting from focusing the light.



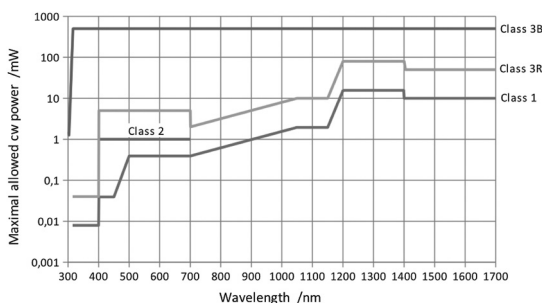
It is rare for lidars to be operated below $0.4\ \mu\text{m}$ for **eye safety**, but it is common for them to be operated at $1.5\ \mu\text{m}$, or longer. A slight decrease in the maximum allowed flux levels occurs as the wavelengths become longer than $1.5\ \mu\text{m}$. This is because light absorption by water in the eye is higher at the surface at longer wavelengths. At $1.5\ \mu\text{m}$, the light is absorbed before it hits the retina, but it is absorbed in a volume of most of the eye. At $10\ \mu\text{m}$, water is more absorptive, so it is absorbed near the surface of the eye. The

chart shows the **maximum permissible exposure** [MPE, given in energy per area (J/cm^2)] for various wavelengths. This level of exposure in 10 sec is often used as a threshold.



Laser Safety Categories

Lasers have been classified using wavelength and maximum output power into four classes and a few subclasses since the early 1970s. The classifications categorize lasers according to their ability to damage human tissue. There are two classification systems, the “old system,” which was used before 2002, and the “revised system,” which started being phased in after that. The latter reflects the greater knowledge of lasers that has been accumulated since the original classification system was devised. It permits certain types of lasers to be recognized as having a lower hazard than was implied by their placement in the original classification system.



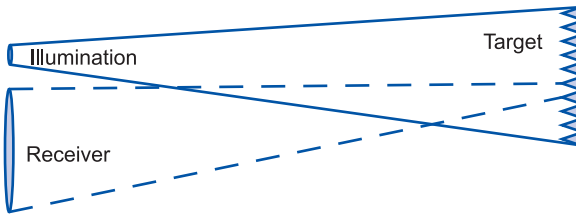
- **Class 1** lasers are safe unless focused.
- **Class 2** covers the spectral range of 400–700 nm. These lasers are safe because of the **blink reflex**, which applies only to visible spectral regions because you if you see it, you will blink. They are 1 mW (cw) or less.
- **Class 3** lasers are ≤ 5 mW cw and are safe if handled properly.
- **Class 3B** lasers might need goggles and can be up to 30 mW.
- **Class 4** lasers can burn the skin and damage the eye.

Caution: Some laser pointers are more than 5 mW. Many of those available for purchase on the Internet are higher than 5 mW.

For a lidar, the **MPE** values on the page 10 are more useful because they are given in energy per area of aperture.

Monostatic versus Bistatic Lidar

One way to isolate the transmit and receive functions is to use a different telescope for each function. Then backscatter is eliminated because the only backscatter of concern comes from scattering off of the aerosols close to the lidar, and the two apertures do not even need to be close to each other. If the transmit and receive apertures are two separate apertures, the setup is a **bistatic lidar**.



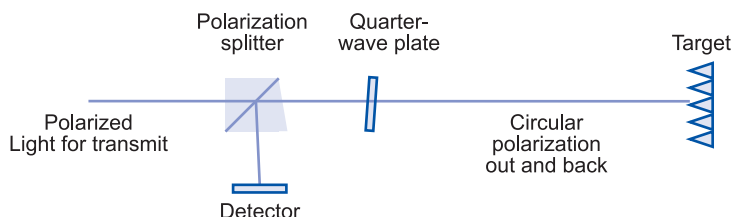
The **transmit aperture**, or **illuminator aperture**, does not have to be the same size as the **receive aperture**, or in the same position. A laser designator is a bistatic lidar illuminator with the seeker as the receiver, even though engineers do not normally think of a laser designator system as being a bistatic lidar.

One reason for using **monostatic lidars** is to save weight and space by not having a second aperture for illumination; however, with bistatic illumination, often the illumination aperture is much smaller than the receive aperture, reducing the size, weight, and power impact of having two apertures. This is the case for **flash lidar**, where the illuminated area is larger than the area viewed by a single DAS. Having a small transmit aperture for illumination is especially important for coherent lidar, so that the phase across the illuminated region is constant. Also, there are advantages to having more than one transmitter and more than one receiver, as in **multiple input, multiple output (MIMO) lidars** (see pages 59–60). Any MIMO lidar is inherently bistatic.

Transmit/Receive Isolation

High-power backscattered laser radiation can blind or damage the receiver, so the high-power transmitted laser waveform needs to be isolated from the receiver. One way to isolate the **receiver** from the **transmitter** is to use a bistatic lidar with separate apertures. If the lidar has a pulsed (or low duty cycle) transmitter, another way to isolate the receiver from the transmitter is to keep the receiver off while the laser is transmitted. With the receiver off, there is no concern about the level of isolation between the transmitter and the receiver.

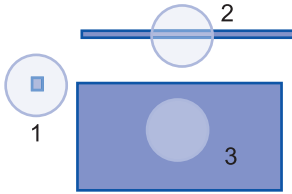
However, there are reasons that a high duty cycle, or cw waveform might be desirable. In this case we need to prevent emitted laser power from being backscattered into the receiver. The most common way to do this is to use **polarization**. If one transmits a linearly polarized laser beam, the beam can be isolated using a polarizing beamsplitter. A quarter-wave plate will convert linear polarization to circular polarization.



When that circular polarization bounces off an object, most of the light will reverse its handedness. On return, the opposite-handed circular polarization beam will be converted to the opposite linear polarization of the laser and will be transmitted to the **polarizing beamsplitter**, and then to the receiver. With careful design, this polarization method of **transmit/receive isolation** can achieve up to 40–45 dB of isolation. The backscatter to hit the camera will come from the quarter-wave plate and the transmit telescope. Quality engineering will reduce this scatter.

Lidar Range Equation

The **radiance** (W/cm^2) for the simple flat-top beam is the laser power divided by the area of the beam footprint, which is a significant antenna gain compared to radiating in all directions. Cross section is not a physical area but is related to the physical area. It represents the percentage of reflected light from a given area of the target.



The figure shows (1) a point target much smaller than the illuminating beam, (2) a linear target larger than the illumination beam in one dimension, and (3) an area target larger than the illumination beam in both directions.

Cross section is constant with range for the last case, but varies with range, or range squared, for the first two cases. Some **lidar range equations** use **target reflectivity** rather than cross section and, therefore, can have R^3 or R^4 in the denominator. A larger receiver aperture then captures more light. Losses occur in the lidar system and in the two-way atmospheric path to and from the target. The total **optical power received** at the detector is therefore given by

$$P_S = P_T \cdot \frac{\sigma}{A_{\text{illum}}} \cdot \frac{A_{\text{rec}}}{\pi R^2} \cdot \eta_{\text{atm}}^2 \cdot \eta_{\text{sys}}$$

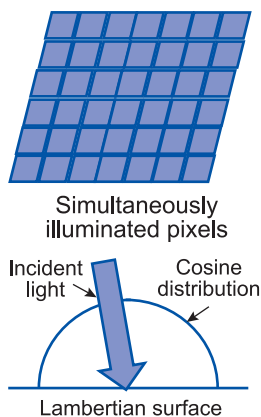
where P_S is the signal power received, P_T is the power transmitted, σ is the cross section in square meters, A_{illum} is the area illuminated, A_{rec} is the area of the receiver, R is the range, η_{atm} is the transmission efficiency through the atmosphere, and η_{sys} is the receiver system optical efficiency. The power received is the power transmitted times two ratios of areas, times appropriate efficiency terms. The first ratio of areas is the cross section divided by the illuminated area at the object plane. The second ratio of areas is the receiver aperture area divided by the effective average area illuminated by Lambertian reflection. This is for the EO-based cross-section definition discussed on page 15.

Lidar Cross Section

For a target with an illumination area larger than the DAS, we can assume that the cross section σ viewed by a receiver pixel is the projected area of a pixel times the reflectance. For square receiver pixels, the cross section is

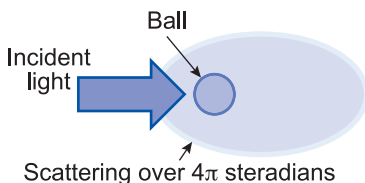
$$\sigma = \rho_t A_p = \rho_t d^2$$

where d is the length of one side of the projected area of a pixel on the target, ρ_t is the reflectance of the area, and A_p is the area of the pixel. Higher spatial resolution means that each detector sees a smaller area and cross section. Surfaces with higher reflectivity have a higher cross section. **Corner cubes** have a much higher cross section because light reflected from them is returned in a small angular cone. The accepted definition of cross section for lidars is different from that for microwave radars. For **lidar cross section**, it is usually assumed that scattering is near Lambertian, and reflected light is reflected into π steradians. We arrive at π steradians as the effective solid angle of reflected light by assuming a cosine distribution of



reflected light over a hemisphere (2π steradians) for Lambertian scattering of light from a rough surface. In microwave radar, the cross-section definition usually involves scattering light over 4π steradians from a small round gold ball.

This makes sense for radar, where often the radar wavelength is longer than the size of the diameter of the round gold ball.

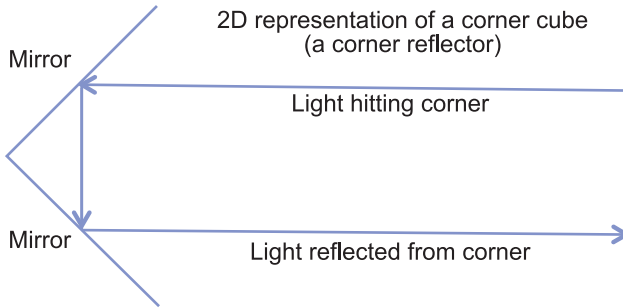


For the EO cross section, the definition is less applicable because the ball is much larger than the wavelength and would block forward radiation. The cross section in either case is proportional to both the area of the target

being illuminated inside one DAS and the surface reflectivity.

Cross Section of a Corner Cube

The **cross section of a corner cube** is much larger than that of other surfaces because light hitting the corner cube is reflected back toward the source. It is like an antenna with gain. A 2D representation of a corner cube is shown.



Light hits the corner cube, bounces twice, then returns toward the exact place the light originated. This is the basic theory of a corner cube. If everything worked as just described, the returned light would be a delta function in angle. However, because a corner cube is limited in size, diffraction will limit the width of the return beam.

As stated on page 15, a Lambertian scatterer has an average **solid angle return** of π . Now we have a solid angle return that is about λ/a in each dimension. The increase in effective cross section compared to a flat plate is the ratio of these solid angles. In one case, the return is spread over a wide angle, whereas a corner cube concentrates the return beam, making the radiance of the return beam larger.

A corner cube can be a **dihedral**, which has two planes meeting and concentrates light in only one dimension. Alternatively, a corner cube can be a **trihedral**, concentrating light in two dimensions. It is generally assumed that a corner cube is a trihedral, in which case, the corner cube cross section is

$$\sigma \approx \frac{4a^4}{\lambda^2}$$

Speckle

When an object is illuminated with a narrowband laser beam, light and dark areas are seen in what is called **speckle**. Speckle is **interference** that occurs when scattering from one part of an object interferes with scattering from another part of the object. Because interference averages out the many different wavelengths in the illuminating light, speckle is not seen when sunlight illuminates an object. Using a laser with a broad spectral band eliminates speckle for a direct detection laser for the same reason that sunlight does not have speckle.



A new **speckle realization** requires a 2π phase difference between the two wavelengths illuminating a target. Speckle depends on surface roughness Δz and coherence length. If instead of a rough surface, we had only two distances within a reflected area, the change in wavelength needed to obtain a new speckle realization would be given by

$$\Delta\lambda = \frac{\lambda^2}{\Delta z - \lambda}$$

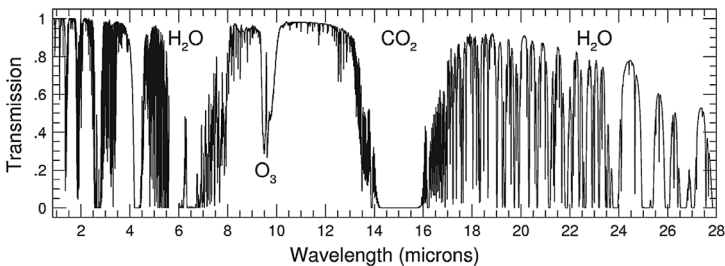
A second way to average out speckle is to change the angle enough to obtain different interference and, thus, different speckle.

The size of an illuminated object determines the extent of speckle lobes, just as the size of an aperture determines the diffraction limit (of diffraction from the scattered return). Interference can occur only over the illuminated area of an object. This means that an object's size can be determined by measuring the angular extent of speckle lobes. The smallest area of interference is set by the sensor resolution and is limited either by sampling or by diffraction from the sensor aperture.

Atmospheric Absorption

The atmosphere has **nitrogen**, **oxygen**, **CO₂**, **water vapor**, and many other smaller components. All of these components can absorb light. Unlike atmospheric scattering, **atmospheric absorption** tends to be very wavelength specific for a given molecule, so small shifts in laser wavelength can make a significant difference in the absorption component to the exponential decay of laser radiation while traversing the atmosphere.

Popular **lidar wavelengths** have traditionally been 1.064 μm because of Nd:YAG availability and 1.55 μm because of its eye safety and the availability of telecommunications components. While absorption is not zero at these wavelengths, both of these regions avoid any significant atmospheric absorption. Recently, there has also been some interest in $\sim 2\text{-}\mu\text{m}$ wavelength and in the MWIR region. Atmospheric absorption at 2 μm and in the MWIR is also not significant, but scattering is less at longer wavelengths. Real beam resolution, of course, is better at the shorter wavelengths.



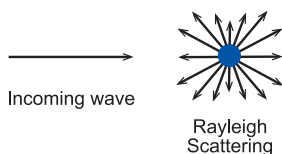
When selecting a lidar wavelength, the designer chooses one of the atmospheric lines that has low absorption. Many of those lines are spread across a fairly wide range from the visible to LWIR. Note that the graph does not assume a large presence of water vapor in the air. At high absolute humidity, the LWIR wavelength region does have significant absorption so should be avoided for those conditions. However, many other parameters will set the exact wavelength chosen.

Atmospheric Scattering

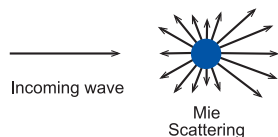
One of the main loss factors in the lidar equation is **atmospheric loss**, much of which is due to scattering. The three main regimes of scattering are:

1. where the particle is much smaller than the electromagnetic (EM) wavelength,
2. where the particle is on the order of the size of the EM wavelength, and
3. where the particle is much larger than the EM wavelength.

Rayleigh scattering is the elastic scattering of light by molecules and particulate matter much smaller than the wavelength of the incident light. Rayleigh scattering intensity is proportional to the sixth power of the diameter of the scattering particle and is inversely proportional to the fourth power of the wavelength of light. This means that shorter wavelengths in visible light (violet and blue) are scattered more strongly than the longer wavelengths toward the red end of the visible spectrum.



Mie scattering theory is used for particles on the order of the same size as the wavelength of the EM radiation. Mie scattering has more forward scatter than scatter in other directions. Particles much larger than the wavelength of light have mostly forward scattering. For any regime of scattering or other losses, attenuation in the atmosphere tends to have an exponential decay at a given wavelength. Each wavelength has a different exponential decay (Beer's law):

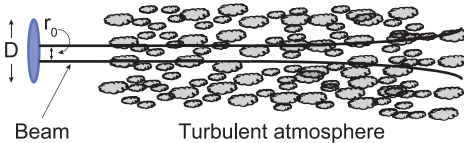


$$E = E_0 e^{-(\beta R)}$$

Scattering is only one component of that decay. Aerosols are one particle to consider, but dust, fog, sand, and other particles can contribute to atmospheric scattering.

Atmospheric Turbulence

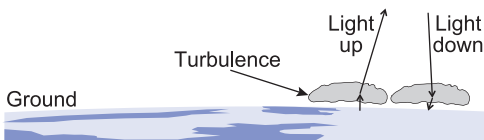
Because of small variations in the **index of refraction** of the air, there is a maximum effective aperture size on receive when a lidar is looking through the atmosphere. This limits the diffraction-limited resolution of a lidar



unless **turbulence** is compensated. While these index of refraction variations are small (on the order of 0.0001 or less), they do cause phase changes across an aperture. The largest aperture diameter possible without corruption due to phase variations is called the **Fried parameter** r_0 .

Because Fried named this parameter r_0 , people sometimes incorrectly call it the Fried radius, but it is a diameter.

When looking through turbulence, the effects are not symmetric. Imaging from the ground looking up through turbulence, the angular deflections are large. While looking



down, turbulence has little effect. The **isoplanatic patch** is the physical

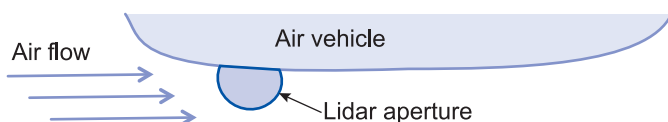
area at any location over which the speckle patterns are the same. The **isoplanatic angle** is the angular region over which the speckle patterns are the same.

The **frozen atmosphere assumption** states that atmospheric turbulence does not change but simply moves from one place to another at a certain velocity, which can be due to wind or vehicle motion. The higher the velocity the faster the atmospheric turbulence changes at a given location. The **atmospheric coherence time** at a given location is given by

$$\tau_0 = \frac{r_0}{v}$$

Aero-optical Effects on Lidar

When a lidar is attached to an air vehicle, a rapid variation in the index of refraction of air near the surface of the vehicle can occur. This variation is due to the **aero-optical effect** and is most pronounced when the lidar aperture protrudes into the wind stream, as shown in the figure.



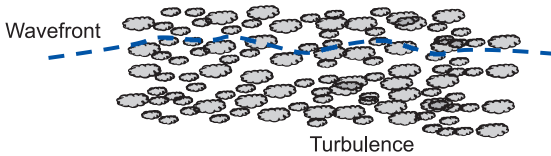
Gimbals are often used to point a lidar. A **gimbal** is a round half-ball that protrudes down into the wind stream, creating additional **turbulence**. Even a smooth surface will create some turbulence. According to the **frozen atmosphere assumption**, coherence time varies with the velocity of the wind or vehicle motion. In this case, the important velocity is that of the vehicle traveling through the air, which can be much faster than wind velocity.

It is not surprising that aero-optical effects contain much higher frequency variations than normal turbulence. The frozen atmosphere assumption is valid until the air vehicle disturbs the air it is flying through, creating turbulence. The aero-optical effect applies in velocities ranging from 50 m/sec to ~mach 1 (~300 m/sec), and some vehicles can fly at supersonic speeds, creating a significant disturbance in the air.

When passively sensing through an **aero-optical layer**, the index variations are close to the air vehicle, so they are essentially in the pupil plane of the sensor. However, when emitting a laser meant to illuminate an object, the beam travels through the aero-optical turbulent layer immediately after being emitted, so that layer is positioned to cause a significant impact on the illuminator beam. Fortunately, for many lidars, the illuminator beam can be fairly large, simultaneously illuminating a large area to be imaged by an array of detectors. This can mitigate the aero-optical effects on the outgoing beam because this beam does not need to be small.

Extended (Deep) Turbulence

It is relatively easy to compensate for turbulence in the pupil plane of a sensor. Unfortunately, most turbulence tends to be distributed between the target and the sensor. Turbulence in the pupil plane creates a **phase change**, as does an **adaptive optics mirror** in the pupil plane, so an adaptive optics mirror with opposite phase change adds to the pupil plane atmospheric phase change to cause zero effective phase change.



When the turbulence is spread throughout the laser path length, phase compensation alone does not completely counteract the turbulence because as light travels from the far field to the pupil plane, some phase information is converted to amplitude information.

In theory, if an adaptive optics aperture could impose plus/minus phase and amplitude variations, extended turbulence could be exactly compensated. If the only amplitude variation that can be imposed is a reduction in amplitude, this is a loss. So, a **deformable mirror** (DM), even with amplitude reduction added, cannot completely compensate for deep turbulence. While a DM cannot compensate for **amplitude variations** in the pupil plane, a number of engineers have shown that the major effect on an image is the phase, not the amplitude. Still, **phase correction** alone is not a complete correction.

A second way to compensate for extended turbulence is to use a computer to insert a **virtual phase screen** at multiple locations in the path to compensate for the real turbulence that occurs at multiple locations.

If the turbulence is not too strong, the lack of precision in compensating for extended turbulence is not a significant issue. Thus, adaptive optics will compensate for mild extended turbulence in the atmosphere.

Laser Power for Lidar

Multiplying both sides of the **power received equation** by time results in the **energy received equation**, which can then be solved for the **required transmitted energy**:

$$E_T = E_R \cdot \frac{A_{\text{illum}}}{\sigma} \cdot \frac{\pi R^2}{A_{\text{rec}}} \cdot \frac{1}{\eta_{\text{atm}}^2} \cdot \frac{1}{\eta_{\text{sys}}}$$

In this equation it is possible to replace area illuminated with range squared times antenna gain, the gain compared to some reference such as π steradians. The power and energy equations presented so far only calculate how much power (or energy) is received back from a given amount of transmitted power (or energy). This says nothing about how much energy per waveform is needed to detect an object. Thus, for various types of lidars, it is necessary to determine the amount of energy required to achieve a certain combination of probability of detection and probability of false alarm. If a higher probability of false alarm is allowed, a higher probability of detection for a given energy detected can be achieved.

A comparison of the returned signal to multiple noise sources is needed for **direct detection**. If the returned signal is amplified, additional noise is introduced, but the overall **signal-to-noise ratio** (SNR) is usually improved.

The two forms of amplified direct detection lidar are **linear-mode avalanche photodiode** (LMAPD), and **Geiger-mode avalanche photodiode** (GMAPD). GMAPDs are often used with very low probability of detection per pulse, requiring many pulses to achieve the required detection probability.

A GMAPD lidar will often require 50–100 measurements to achieve the required probability of detection and false alarm.

To mitigate noise effects in **heterodyne detection**, instead of amplifying the signal, the power of the LO is increased. However, shot noise depends on the total signal detected, so the LO contributes to additional shot noise.

Lidar Signal-to-Noise Ratio

Next to be addressed is how much energy per waveform is required to make a detection. To achieve detection, a certain **SNR** is required:

$$SNR = \frac{\langle i_s^2 \rangle}{\langle i_n^2 \rangle}$$

This equation expresses a ratio of the mean-squared signal current to the mean-squared noise current equaling some value. That value can vary depending on circumstances. The laser power can be increased to increase SNR. However, it is often less expensive to mitigate mean-squared noise power. This can be done for both **direct detection lidar** and **coherent lidar** by using gain early in the detection process so that any noise after the gain is not as influential.

The SNR for coherent lidar can also be increased by increasing the power of the LO. Because of the option of increasing **LO power** to mitigate noise effects, pre-amplification gain has not been used as often in coherent lidar as in direct detection lidar.

The mean-squared **signal current** is

$$\langle i_s^2 \rangle = \begin{cases} G^2 \Re^2 P_s^2 & \text{(direct detection)} \\ 2\eta_h G^2 \Re^2 P_{LO} P_s & \text{(coherent detection)} \end{cases}$$

where G is pre-amplifier gain, $\Re$ is detector responsivity, P_s is signal power, P_{LO} is LO power, η_h is heterodyne mixing efficiency, and

$$\Re = \frac{\eta_q e}{h\nu}$$

where η_q is detector quantum efficiency, e is the charge on an electron, h is Planck's constant, and ν is the carrier frequency of the light.

The mean-squared **noise current** is

$$\langle i_n^2 \rangle = \langle i_{\text{shot,sig}}^2 \rangle + \langle i_{\text{shot,LO}}^2 \rangle + \langle i_{\text{bk}}^2 \rangle + \langle i_{\text{dk}}^2 \rangle + \langle i_{\text{th}}^2 \rangle$$

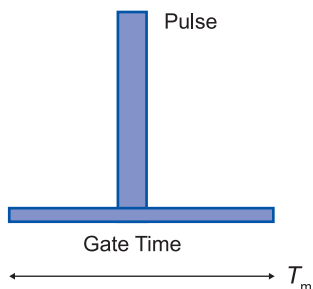
where i_{bk} is background current, i_{dk} is dark current, and i_{th} is thermal noise current. Signal and LO shot noise terms are separated because direct detection does not have LO shot noise.

Direct Detection Signal-to-Noise Ratio

Received optical power converts to current, which must be squared for **direct detection** to calculate the power term used in the **signal-to-noise equation**. The **received optical power** is related to the rate of arrival of the received photons by

$$P_s = \frac{Nhc}{\lambda T_m}$$

where T_m is the time period over which the measurement is made, and N is the number of photons per pixel received during that measurement time. For pulsed lidar systems, T_m will be the receiver gate time. The laser pulse width can be narrower than the gate time, as shown.



When optical signals hit the detector, they create variations in the currents in the receiver electronics; these variations have associated mean levels and noise fluctuations. When the signal and noise sources are considered, the resulting **electrical SNR** is given by

$$SNR = \frac{\langle i_s^2 \rangle}{\langle i_n^2 \rangle} = \frac{G^2 \Re^2 P_s^2}{\langle i_{\text{shot,sig}}^2 \rangle + \langle i_{\text{shot,LO}}^2 \rangle + \langle i_{\text{bk}}^2 \rangle + \langle i_{\text{dk}}^2 \rangle + \langle i_{\text{th}}^2 \rangle}$$

The five terms in the denominator represent the signal shot noise, the LO shot noise (which is not present for direct detection), the background current noise, the dark current noise, and the thermal (or Johnson) noise, respectively. Each of these noise sources will be discussed on subsequent pages.

Noise Probability Density Functions

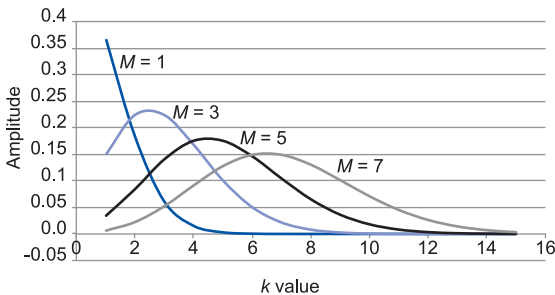
Common **noise probability density functions** we might encounter include Gaussian, Poisson, and negative binomial distributions. A Poisson distribution gives the probability of exactly x events occurring during a time period if the events take place independently at a constant rate. This is a common type of occurrence; for example, dark current would meet this criteria. A **discrete Poisson probability density function** for the distribution of k photons is given by

$$q(k, M) = \frac{M^k e^{-M}}{k!}$$

where M is the number of detections. For a large M , the Poisson distribution approaches a Gaussian distribution. This is called the **central limit theorem**. The **Gaussian probability density function** is

$$p(k) \approx \frac{1}{\sqrt{2\pi M}} e^{-(k-M)^2/2M^2}$$

As M grows larger, the probability density function moves toward the middle and becomes symmetric, or Gaussian. The graph shows that even for $M = 7$, the shape of the Poisson distribution is approaching Gaussian.



A **negative binominal distribution** occurs if events occur as a random variable that is represented by a gamma distribution. A diffuse, or rough, target results in a **gamma distribution** for the integrated intensity at the receiver.

Thermal Noise

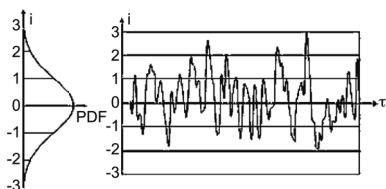
Thermal noise, also known as **Johnson noise** or **Nyquist noise**, is electronic noise generated by the thermal agitation of charge carriers (usually electrons) inside an electrical conductor. This noise occurs regardless of applied voltage. Thermal noise is given by

$$E_{\text{th}} = kTB$$

where k is the Boltzmann constant, T is the temperature in Kelvin, and B is the bandwidth of the detector. High detector bandwidth captures more noise than low detector bandwidth. Higher temperature creates more thermal noise. k is $1.3806488 \times 10^{-23} \text{ m}^2 \text{ kg sec}^{-2} \text{ K}^{-1}$. T is often assumed to be 300 K. The mean-squared thermal noise current is

$$\langle i_{\text{th}}^2 \rangle = \frac{4kTB}{R_L}$$

where R_L is load resistance. Thermal noise is not affected by pre-amplifier gain; therefore, it is one of the noises mitigated by gain. All values above absolute zero have thermal fluctuations.



The plot shows random thermal noise spectral distribution.

Thermal noise is very important in radar, but less important in lidar because visible photons contain more energy than microwave photons, emphasizing the importance of a competing noise called **shot noise**, which consists of additional current fluctuations that occur when a voltage is applied and a macroscopic current starts to flow. The radio frequency community tends to measure thermal noise in decibels above a milliwatt (dBm):

$$P_{\text{thdbm}} = 10\log_{10}(1000 kTB)$$

where 1000 is included because thermal noise is expressed in milliwatts rather than watts.

Shot Noise

Shot noise results from the quantization of an EM wave. Short wavelengths (high frequencies) have more energy per photon because

$$E_p = h\nu$$

is the energy in a **photon**. In optics, wavelength is used more often than frequency, so

$$E_p = \frac{hc}{\lambda}$$

This same quantization phenomenon occurs for microwave frequencies, but a joule of energy has many more photons at these frequencies. At 1.5 μm , a single photon has 1.3×10^{-19} J versus 6.6×10^{-24} J at 10 GHz (3-cm wavelength). A small variation in the arrival rate of photons provides shot noise.

For large numbers, the Poisson distribution approaches a normal distribution, typically making shot noise in actual observations indistinguishable from Gaussian noise except when the elementary events (photons, electrons, etc.) are so few that they are individually observed. Since the standard deviation of shot noise equals the square root of the average number of events M , the SNR is given by

$$SNR = \frac{M}{\sqrt{M}} = \sqrt{M}$$

Thus, when M is very large, the SNR is very large as well, and any relative fluctuations in M due to other sources are more likely to dominate over shot noise. However, when the other noise source is at a fixed level, as in the case of thermal noise, increasing M (the DC current or light level, etc.) can sometimes lead to dominance of shot noise. The shot noise due to the signal return is

$$\langle i_{\text{shot,sig}}^2 \rangle = 2eG^2FRP_sB$$

where F is the excess noise factor associated with the pre-amplifier gain. For **coherent lidar** with a strong **LO**, the shot noise associated with the LO (identical in form to the signal shot noise equation) is often a dominant noise source.

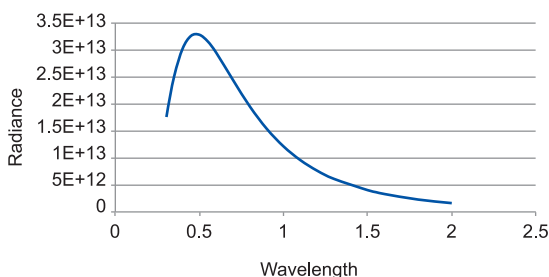
The Sun as Background Noise

Background noise is just like signal, but it is unwanted signal. In the visible wavelengths, the sun is an ideal example of background noise. Background noise can often be reduced by using a narrowband filter and a narrowband laser. The sun is approximately a 6050 K blackbody. The diameter of the sun is 1.39×10^9 m, or 9.6×10^{-3} rad as viewed from the earth, and it is 1.496×10^{11} m from the earth. The sun provides about 1.36 KW/m² above the atmosphere. Narrower filters reject more sunlight.

The **radiance per wavelength** is

$$L_{\lambda} = \frac{2hc^2}{\lambda^5} \left[\exp\left(\frac{hc}{\lambda kT}\right) - 1 \right]^{-1}$$

where a very strong dependence on wavelength is evident, c is the speed of light ($\sim 3 \times 10^8$ m/sec), and k is the Boltzmann constant (1.38×10^{-23} J/K).



The spectral trends of sun radiance are evident, assuming a 6050 K blackbody for the sun. The peak radiance of the sun is at $0.5 \mu\text{m}$, and $1.55 \mu\text{m}$ has about $0.15 \times$ peak sun radiance. To calculate the background radiation from the sun, the sun needs to reflect off of the area being viewed by a pixel, usually using **Lambertian scattering**. The reflected radiation of the sun provides background radiation. Solar glints can provide higher spectral background radiation. The background noise is

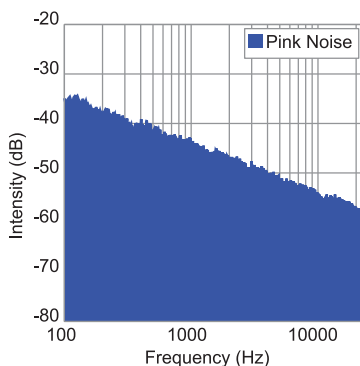
$$\langle i_{\text{bk}}^2 \rangle = 2eG^2 F \mathcal{R} P_{\text{bk}} B$$

which is identical to the signal shot noise term, as background noise is a form of signal.

Dark Current, $1/f$, and Excess Noise

Dark current noise is the noise in a detector that is not illuminated. The charge generation rate in a detector can be related to specific crystal defects within the depletion region of a detector (bulk defects). Bulk dark current participates in pre-amplifier gain. **Bulk defects** are the most likely sources for dark current. Surface defects can also contribute to dark current, but are much less likely to do so. Surface dark current does not participate in pre-amplifier gain. Both bulk- and surface-generated dark current primary carriers are Poisson distributed. The pattern of different dark currents can result in a **fixed-pattern noise**. **Dark frame subtraction** can remove an estimate of the mean fixed pattern, but a temporal noise remains because the dark current itself has a shot noise.

Dark current is not noise, but unwanted signal. However, like any signal, it does *contain* noise (fluctuations). When a detector is biased, even before it is illuminated, some noise “signal” occurs. If the effects of dark current and other noise are mitigated by amplification, **excess noise** can be associated with the amplification.



The noise from dark current is

$$\langle i_{\text{dk}}^2 \rangle = 2eB(I_{\text{dks}} + I_{\text{dkb}}G^2F)$$

where I_{dks} is surface dark current that is not multiplied in an **APD**, and I_{dkb} is bulk dark current that is multiplied by the gain of an APD. When a signal is amplified, noise is added, shown by the

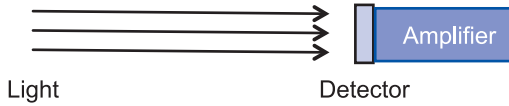
excess noise factor. **$1/f$ noise** (also called **flicker noise**) is a signal or process with a frequency spectrum such that the power spectral density (energy or power per hertz) is inversely proportional to the frequency. $1/f$ noise is not well understood and is sometimes called **pink noise**.

Avalanche Photodiodes and Direct Detection

One way to reduce the effect of noise in **direct detection** is to use gain. This gain can be from a fiber amplifier before detection, but the most common type of gain is **avalanche gain** after detection. Expanding the noise terms results in

$$SNR = \frac{\langle i_s^2 \rangle}{\langle i_n^2 \rangle} = \frac{G^2 \Re^2 P_s^2}{2eBG^2F[\Re(P_s + P_{LO} + P_{bk}) + P_{dks}] + 2eBi_{dks} + 4(kTB/R_L)}$$

Gain after detection reduces the influence of surface dark current and thermal noise. Background radiation is amplified in the same way that the signal is amplified. Surface dark current can also be amplified in an **APD**.



Amplification adds noise. For a background-limited direct detection receiver,

$$SNR = \frac{\eta_D P_s^2}{2h\nu B P_{bk}}$$

Within the limit in which the signal shot noise dominates the other noise terms, the SNR is given by

$$SNR_{\text{shot,lim}} = \frac{\eta_D P_R}{2h\nu B}$$

Within this limit, the SNR is directly proportional to the number of photons received. This is the best that can be achieved and is the goal of using gain.

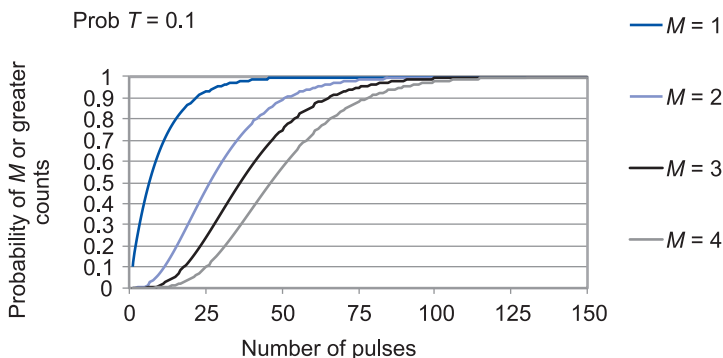
Number of Photons Required for a GMAPD Lidar Camera

Four possible **GMAPD detection events** can trigger the receiver:

1. the detection of a desired target photon,
2. the detection of an undesired foreground clutter photon (such as from backscatter from fog),
3. the detection of undesired background radiation (such as the sun), or
4. the undesired detection of a dark electron.

If many laser pulses are sent out, **coincident returns** (returns in the same range bin) will result for reflection from a target or from fixed foreground objects. Fog, snow, and rain provide distributed returns over their extent. If the fog or clouds are close enough to the target when an avalanche occurs, the dead time will prevent avalanching on a return from the target.

One problem with GMAPDs is that once the avalanche fires, there is a dead time of at least $\sim 500 \mu\text{sec}$, during which the GMAPD cannot be armed without a high a probability of **afterpulsing** (unless the arm times are very short, thus reducing the chance of an afterpulse).

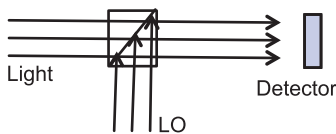


The probability of detecting a specified **number of photons** M backscattered from the object of interest out of N pulses can be described by the binomial distribution as follows:

$$P(M \text{ out of } N \text{ pulses}) = \frac{N!}{M!(N-M)!} (P_o)^M \times (1 - P_o)^{N-M}$$

Heterodyne Detection

When the return signal interferes with an **LO**, the resulting optical intensity can fluctuate in time or in space. For **spatial heterodyne detection**, also called **digital holography**, the LO is at the same frequency as the signal but is offset to create spatial fringes. For **temporal heterodyne detection**, the LO



and return signal are aligned as shown. Fluctuations occur near DC and at the difference and sum frequencies of the two fields. For optical frequencies,

the higher-frequency power oscillations are well beyond the maximum frequency response of detectors. The only detectable power fluctuations are near DC and at the difference frequency. The signal and LO electric fields and the heterodyne detection current are:

$$\begin{aligned}\vec{E}_{\text{sig}} &= E_{\text{sig}} \exp(k_{\text{sig}} \vec{r}_{\text{sig}} - j\omega_{\text{sig}}t) \\ \vec{E}_{\text{LO}} &= E_{\text{LO}} \exp(k_{\text{LO}} \vec{r}_{\text{LO}} - j\omega_{\text{LO}}t) \\ I &= \vec{E}_{\text{sig}} \cdot \widehat{\vec{E}}_{\text{sig}} + \vec{E}_{\text{LO}} \cdot \widehat{\vec{E}}_{\text{LO}} + \vec{E}_{\text{sig}} \cdot \widehat{\vec{E}}_{\text{LO}} + \vec{E}_{\text{LO}} \cdot \widehat{\vec{E}}_{\text{sig}}\end{aligned}$$

The electrical power measured in a heterodyne receiver is linearly proportional to the optical power received because the terms with signal $\times$ LO power are at the intermediate frequency (IF). The factor of two is eliminated from the denominator due to the same factor in the signal $\times$ LO power. The LO adds shot noise, which is accounted for in the denominator:

$$\begin{aligned}SNR &= \frac{\langle i_s^2 \rangle}{\langle i_n^2 \rangle} \\ &= \frac{\eta_h G^2 \Re^2 P_{\text{LO}} P_s}{eBG^2 F [\Re(P_s + P_{\text{LO}} + P_{\text{bk}}) + P_{\text{dks}}] + 2eBi_{\text{dks}} + 4(kTB/R_L)}\end{aligned}$$

The LO terms are added to the SNR equation. The LO can be increased to dominate other noises. Engineers have recently used APDs with heterodyne detection, eliminating the SNR-generated need for a strong LO. The ideal SNR for heterodyne detection is given by

$$SNR = \frac{\eta_D \eta_h P_s}{h\nu B}$$

Temporal Heterodyne Detection

For **temporal heterodyne detection**, only the time terms of the heterodyne equations are kept. Eliminating the spatial terms,

$$\vec{E}_{\text{sig}} = E_{\text{sig}} \exp(k_{\text{sig}} \vec{r}_{\text{sig}} - j\omega_{\text{sig}} t)$$

becomes

$$\vec{E}_{\text{sig}} = E_{\text{sig}} \exp(-j\omega_{\text{sig}} t)$$

and

$$\vec{E}_{\text{LO}} = E_{\text{LO}} \exp(k_{\text{LO}} \vec{r}_{\text{LO}} - j\omega_{\text{LO}} t)$$

becomes

$$\vec{E}_{\text{LO}} = E_{\text{LO}} \exp(-j\omega_{\text{LO}} t)$$

Therefore, when the signal and LO interfere, the following results:

$$I = 2E_{\text{sig}} \cdot E_{\text{LO}} \cdot \exp[-jt(\omega_{\text{sig}} - \omega_{\text{LO}})]$$

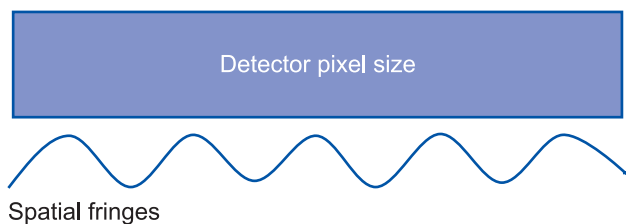
The LO is aligned with the return signal. The LO frequency is generally offset more than any Doppler shift expected in the return signal. For the well-designed **heterodyne detection receiver** (with sufficient LO power), the SNR is proportional to the number of photons received, even when the signal is very weak; whereas, for the **direct detection receiver**, the signal hitting the detector must be strong enough that its shot (photon) noise dominates all other noise.

If a **coherent receiver** is truly shot-noise limited, and if the heterodyne efficiency is unity, the coherent SNR will always be greater than or equal to the direct detection SNR (assuming detectors having the same quantum efficiency).

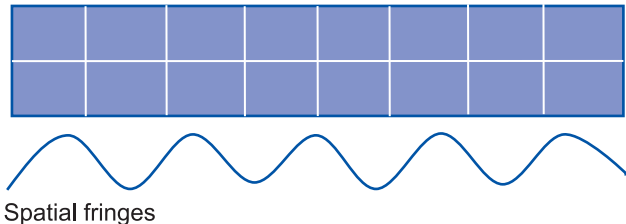
This does *not* mean that the probability of detection and the probability of false alarm are always better for coherent detection, as those probabilities depend on the statistical fluctuations of the signal and noise (primarily signal).

Heterodyne Mixing Efficiency

In order to retain only the time portions of the signal and LO, the signal and the LO must overlap. If perfect overlap does not occur (and it never does), some loss occurs, which is called **heterodyne mixing efficiency**. The quadratic phase terms in the LO and the signal beams need to match; otherwise, the phases will interfere to create rapidly varying **spatial fringes** that will be difficult to sample.



As seen in the figures, if the detector is much larger than the fringe spacing, many fringes are contained in one detector, and the average value becomes close to zero.



With finer detector spatial sampling, or fringes with lower spatial-frequency content, peaks and valleys will occur across the detector array. When using **temporal heterodyne detection**, attention is paid to the temporal portion of the beat between the LO and signal, so these spatial fringes are, at best, a distraction and, at worst, they zero out the beat between the LO and the signal.

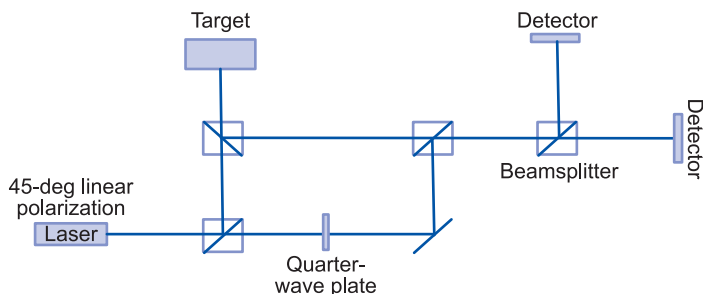
A careful design of the LO is needed such that it illuminates the detector array in a manner that matches the return signal. This can be at a focus or with uniform illumination, but the key is to match the LO with the signal return. A powerful technique to achieve this is to back propagate the LO through the receive optics and the signal transmission path to determine what shape the LO should have to provide an exact signal match.

Quadrature Detection

In **quadrature detection**, the **intermediate frequency** signal (the beat frequency between the LO and the signal) is divided into two parts. One portion of the signal is delayed by 90 deg in phase from the other portion. This allows the intensity and the phase portions of the returned signal to be separated.

The most direct way to measure the two components of the complex signal is to direct the signal to two detectors, each with its own **reference beam**. If the two reference beams are the I component and the Q component of the transmitted beam, the two detector outputs will be the real and the imaginary components of the signal, respectively.

This technique is commonly used in radar when the wavelength is long. Short optical wavelengths make it impractical to maintain the phase relationship between the references because path lengths must be held constant to a fraction of a wavelength. However, the circularly polarized reference is mixed with a signal beam that is linearly polarized at 45 deg. This allows for the separation of the real from the imaginary part of the interference term by measuring the intensity in both its horizontally and vertically polarized components.



The figure shows a laser with 45-deg linear polarization. One leg travels through a quarter-wave plate and becomes circularly polarized. The leg is then beat against the 45-deg linearly polarized laser beam. This creates two beams with 90-deg phase difference that can be separated by a polarizing beamsplitter.

Carrier-to-Noise Ratio for Temporal Heterodyne Detection

For heterodyne detection, the conversion from received optical power to electronic power is linear because the surviving term after interference is a term that multiplies the LO by the return signal, thereby having only one factor of returned signal.

In heterodyne detection, the electronic power received is directly proportional to the optical signal power.

Received optical power converts to current, which must be squared in order for direct detection to obtain electronic power as used in signal-to-noise calculations; however, in this case, instead of squaring the terms, the terms are multiplied by the **LO**.

The **carrier-to-noise ratio** (CNR) is the **SNR** on this beat frequency between the signal and the LO. This refers to the mean electrical power at the **IF** carrier frequency (difference frequency) compared to the noise power.

$$I = 2E_{\text{sig}} \cdot E_{\text{LO}} \cdot \exp(k_{\text{sig}} \vec{r}_{\text{sig}} - k_{\text{LO}} \vec{r}_{\text{LO}})$$

The term CNR came from the radio frequency (RF) community, where it is generally accepted to be a predetection measurement, i.e., one made at RF. In cable industry vernacular, SNR is generally accepted to be a premodulation or postdetection measurement, i.e., one made on a baseband signal such as video or audio. For a lidar, SNR is often measured after speckle averaging or another technique that makes the signal more detectable.

The carrier-to-noise ratio is the mean electrical power at the IF carrier frequency (difference frequency) compared to the noise power.

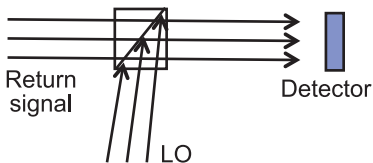
In the RF context, CNR is a predetection measurement performed on RF signals. SNR is a premodulation or postdetection measurement performed on baseband signals. CNR is often given (in decibels) as the ratio of the level of the carrier to that of the noise in a receiver's IF bandwidth before any nonlinear processes such as amplitude limiting and detection takes place.

Spatial Heterodyne Detection / Digital Holography

A major advantage of **spatial heterodyne detection** is the ability to use low-bandwidth framing cameras, which are available in large formats. Visible cameras have very large formats. At $1.5\text{ }\mu\text{m}$, camera formats are smaller, but 512×640 pixels is a very common format, and megapixel cameras are becoming available.

A simple case of a plane wave signal (something viewed at a distance) and a plane wave LO in the pupil plane is considered. The temporal part of the beat frequency does not matter in this case. The LO and the return signals are at the same frequency, except for velocity considerations, which will be initially ignored. Scattering off of a point reflector is assumed, and the LO is emitted at a point the same distance from the aperture as the point reflector but slightly offset in x so that it hits the detector array at an angle that differs from the return angle. The LO is slightly offset in the x direction.

We want to calculate the spatial variation in x of the LO and return signal interfering with each other. The distance from the point reflector to x is defined as x_0 . To obtain exactly the right curvature, the LO is placed the same distance from the center of the imager as the point reflector is; thus, if the LO is emitted as a point radiator slightly to the side, it would



need to be slightly closer in the z direction (this is a minor change that can be ignored). It is assumed that z is significantly larger than the x displacements.

Only one term oscillates in x ; the other term is a constant. The oscillation is the product $x \times x_0$, so the farther from zero we place the LO the more rapid the oscillation in spatial frequency. As we move farther off axis, the fringes oscillate more rapidly. The detector sampling must be sized to see these fringes.

SNR for Spatial Heterodyne Detection

For **spatial heterodyne detection** we keep only the spatial terms of the heterodyne equations. Eliminating the time terms,

$$\vec{E}_{\text{sig}} = E_{\text{sig}} \exp(k_{\text{sig}} \vec{r}_{\text{sig}} - j\omega_{\text{sig}} t)$$

becomes

$$\vec{E}_{\text{sig}} = E_{\text{sig}} \exp(k_{\text{sig}} \vec{r}_{\text{sig}})$$

and

$$\vec{E}_{\text{LO}} = E_{\text{LO}} \exp(k_{\text{LO}} \vec{r}_{\text{LO}} - j\omega_{\text{LO}} t)$$

becomes

$$\vec{E}_{\text{LO}} = E_{\text{LO}} \exp(k_{\text{LO}} \vec{r}_{\text{LO}})$$

Therefore, when the signal and LO interfere, the following results:

$$I = 2E_{\text{sig}} \cdot E_{\text{LO}} \cdot \exp(k_{\text{sig}} \vec{r}_{\text{sig}} - k_{\text{LO}} \vec{r}_{\text{LO}})$$

The electrical power measured in a spatial heterodyne receiver is also linearly proportional to the optical power received because only the terms with signal $\times$ LO are at the intermediate frequency. From a SNR point of view, this is identical to temporal heterodyne detection, so the same **SNR equation** is appropriate:

$$\begin{aligned} \text{SNR} &= \frac{\langle i_s^2 \rangle}{\langle i_n^2 \rangle} \\ &= \frac{\eta_h G^2 \mathfrak{R}^2 P_{\text{LO}} P_s}{eBG^2 F[\mathfrak{R}(P_s + P_{\text{LO}} + P_{\text{bk}}) + P_{\text{dks}}] + 2eBi_{\text{dks}} + 4(kTB/R_L)} \end{aligned}$$

If necessary, the LO can be increased to dominate other noises. Recently, engineers have used APDs with heterodyne detection, eliminating the need to use a strong LO. Once noises have been minimized by either a strong LO or an APD, the ideal SNR for heterodyne detection is given, as before, by

$$\text{SNR} = \frac{\eta_D \eta_h P_s}{h\nu B}$$

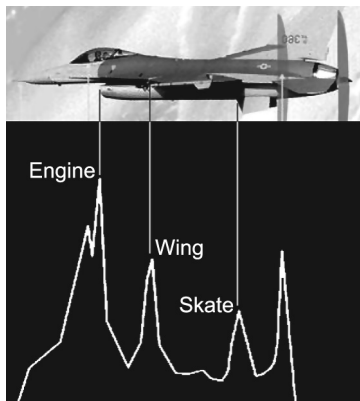
Once again, the SNR is directly proportional to the power received and therefore to the number of received photons.

1D Range-Only Lidar

Range resolution depends only on the **bandwidth of the lidar**. Range can be measured with pulsed waveforms using short pulses or with other waveforms. A 1D image is a range profile of an object. The range profile does not depend on the size of the receiving optical apertures, except for SNR considerations, so it can be used at long range if sufficient SNR is available. This can be very useful in environments that are not very cluttered and for objects at a great distance, where the transverse dimensions of the object are not resolved by the optical system and the only spatial information about the object comes from the **range profile**.

A high-range-resolution profile is most useful when the orientation of the object is known. For example, one generally assumes that an airplane is oriented such that the nose is pointed in the direction the plane is traveling. An airplane's scattering signal varies with range, depending on what scatterers are in a given range bin. Starting from the time when the detector is first triggered by the large scattering signal from the nose, the dimensions of the rest of the airplane can be calculated by the **relative time delay**.

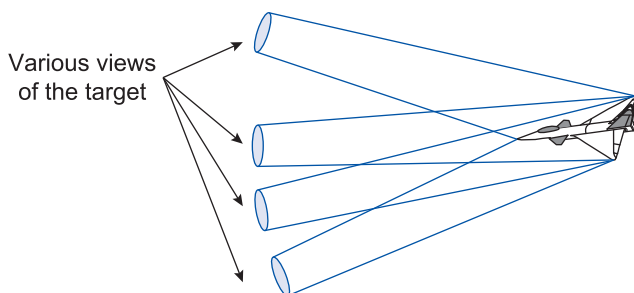
The figure shows the range profile of an airplane viewed in 1D from the front, although the image is shown from the side to allow the reader to correlate the plane with the range profile.



Tomographic Imaging Lidar

If we have multiple high-range-resolution 1D images from different locations, we can develop **angle information** as well as **range information** about the target. The wider the **angular distribution** of the high-range-resolution images the better the angular imaging information. A time-dependent return signal is collected by a single detector to provide a high-range-resolution image. This provides a 1D slice of the 3D spatial profile of the object. When the object rotates or the lidar moves, different slices of the 3D profile can be obtained. If a sufficient number of **1D slices** are collected, a 2D image can be reconstructed.

Various algorithm approaches can be used to combine multiple range profiles to make a more complete image from multiple 1D profile images at different angles. The more complex the object the more necessary it is to obtain range profile information over more and wider angles. Of course, even with a single range profile, 1D imaging can be useful for recognizing some objects, as discussed on page 40.



When a **coherent detection system** is employed and data is collected from multiple views, an object's **Doppler spectrum** can be used to assist in determining the angle from which a given 1D image has been collected. Also, the **rotation speed** of an object can be precisely measured with Doppler techniques. If the object rotates, or the sensor moves, different range projections are formed; thus, knowing the exact rotation speed can make it easier to generate **tomographic images** from multiple 1D images.

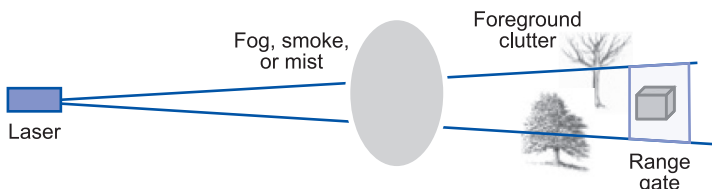
Range-Gated Active Imaging (2D Lidar)

Range-gated active imaging (2D lidar) is similar to passive imaging, but with illumination. A gated framing camera can be used to capture photons in an array. One benefit of this type of lidar compared to passive sensors is that it works at night while using shorter wavelengths. The diffraction limit at shorter wavelengths allows more leeway to use longer-focal-length imaging optics because the diffraction limit is

$$\vartheta \approx 1.03 \frac{\lambda}{D}$$

For the same-sized aperture, higher diffraction-limited resolution can be achieved due to the ability to use a shorter wavelength. Also, because this is active imaging using a laser illuminator, the thermal cross-over problems encountered with thermal imagers are not an issue. To enhance SNR, a **gated 2D camera** is preferred; this camera gathers noise only over a short time period as opposed to continuously gathering noise. This type of lidar frames at low repetition rates, consistent with the repetition rates of framing cameras; 10–30 Hz would be typical.

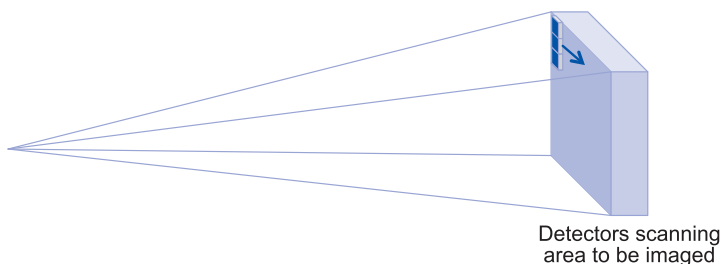
Nanosecond-class laser pulses are not a requirement, but it is likely that **Q-switched lasers** will be used, resulting in 5- to 15-nsec pulse widths. While the pulse widths are likely to be nanosecond class, the detector bandwidth does not need to be that fast. Framing cameras at 30 or 60 Hz can be used, but with the addition of **gating circuitry**, which might be microsecond class. This means that a much simpler read-out circuitry can be used compared to that for most fully 3D-imaging lidar cameras. Another benefit of gated 2D imaging is the ability to gate the obscuring elements that occur before the object of interest, as well as clutter that occurs before or after the obscuring elements.



3D Scanning Lidar

Scanning laser radar is one technology used for collecting 3D images. A single detector, or a small number of individual detectors, is scanned at a rapid rate to make a 3D image. Each detector has sufficiently high bandwidth to measure the time of the return laser pulse accurately enough to provide highly precise range information. This need for a high-bandwidth detector response and read out is the reason that scanning lidars are popular, especially for commercial applications, where cost is a significant issue.

Scanning lidars can obtain detectors more easily and less expensively than flash 3D lidars, but they require higher-speed scanners and improved **stabilization**. A scanning lidar must be stabilized to a small portion of an individual detector, and the scanner must move a small number of detectors over the full image space to create each frame. A scanning lidar is similar to a TV raster scan but creates a 3D image. Two current significant applications for scanning lidars are 3D mapping and driverless cars.



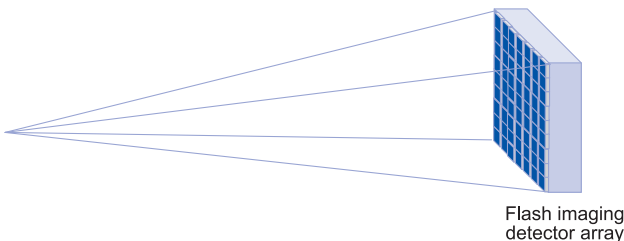
3D scanning lidars have all of the same benefits as 2D lidars but, in addition, provide **3D shape information**. 3D shape is an excellent discriminant for recognizing objects because shape does not change as much as illumination features associated with most 2D imagery. **3D mapping** is becoming widespread because it can be applied in diverse locations such as a downtown area or an area with flood issues.

3D Flash Imaging

Flash imagers use an array of detectors to develop the image. One of the first methods of large-array flash imaging used polarization change to measure range. More recently, high-bandwidth focal plane arrays (FPAs) with appropriate read-out integrated circuits (ROICs) have been developed. In flash imaging:

- The return laser power is imaged onto an array of detectors so the scanning does not need to be so fast, reducing size, weight, and power.
- Stabilization is required only to a fraction of the angular size of the array. This reduction in required stabilization precision can have a significant systems impact.
- The energy returning to the receiver is divided among a large number of detectors. Thus, imaging of targets containing many pixels in angle/angle space requires high peak-illumination power and/or very sensitive detectors compared to single-pixel scanners.

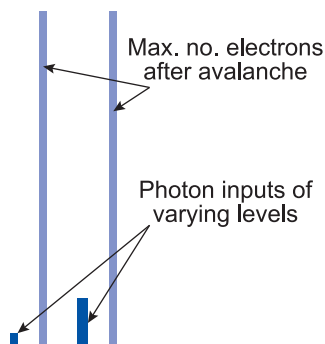
One method to increase the sensitivity of a direct detection receiver is to amplify the return signal so that the signal level is above noise sources introduced downstream from the initial detection. **APDs** are often used to increase sensitivity. LMAPD detectors have a gain that provides a linear relationship between the number of photons received and the amplitude of the signal output from the detector. GMAPDs have a huge gain, regardless of the number of photons input. Fiber pre-amplification before detection can also increase sensitivity but, to date, has not been used with arrays of lidar detectors, although it is used with single detectors for this purpose.



Polarization-encoded imaging addresses a different issue. Instead of using high-bandwidth ROICs, polarization-encoded imagery uses polarization rotation to measure range, meaning that simple framing cameras can be used to perform 3D flash imaging. This topic is further discussed on page 47.

Geiger-Mode APD Flash Lidar

A Geiger-mode receiver temporarily overbiases an **APD** detector for the range gate time during which return signal photons are expected. GMAPDs have a very large avalanche gain for any photon hitting a detector. To avoid afterpulsing, GMAPDs are not armed to detect photons for a period after an avalanche,



called the dead time. They cannot measure the signal intensity or grayscale on each pulse. **GMAPD flash imagers** run at high repetition rates because probability of detection (avalanche) on a single pulse is often purposely kept low. A target is identified by **coincidence processing** (looking for multiple returns coincident in the same range bin), often requiring 50 or 100 laser pulses to obtain 2–4 returns in the same range bin.

Currently, 32×32 or 32×128 GMAPD imagers are available.

A higher reflectance area has a higher probability of return, causing more coincident avalanches to trigger at that range. There is a bias toward nearer ranges because of dead time, but this bias is slight if probability of detection on a single pulse is kept low. This is a major reason that GMAPDs are often run with a probability of detection less than 1 on a single return. With GMAPDs, effective range profile and grayscale both require more photons than simple 3D mapping. For a given range target and transmit and receive aperture diameters, the emitted laser energy needs to be set at the correct level in order to achieve the correct probability of an avalanche.

GMAPD-based flash imaging is a highly efficient method to 3D map an area.

- A key advantage is the relative simplicity of the receiver and data acquisition electronics compared to wide-bandwidth, linear-mode receivers.
- A disadvantage is that when the target has range depth, and/or when grayscale information is needed, the total energy required to map an area is significantly higher than when 3D mapping bare ground.

Linear-Mode APD Flash Lidar

LMAPDs are designed to operate below the breakdown voltage with high gain and low noise. Their operation produces an average photocurrent that is linearly proportional to the incident optical flux (hence, the name *linear mode*). Higher gain can produce more sensitivity compared to lower gain because it more efficiently masks noise generated downstream in the detection process.

LMAPDs in a 128×128 camera format are commercially available with a factor of 10 linear gain. Some LMAPDs can have a factor of up to 1000 linear gain, approaching single-photon-counting capability. LMAPDs made using **HgCdTe** require cooling, even when operating in the NIR region, thus imposing an additional systems issue.

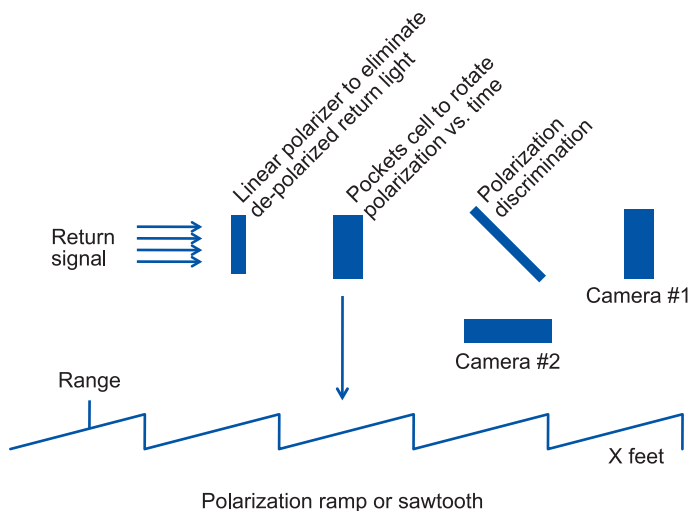
LMAPDs are very useful for shorter-range imaging but less useful for long-range imaging when extreme sensitivity is needed.

Unlike Geiger-mode systems, which must geo-register and aggregate multiple high-frame-rate returns, linear-mode systems have the advantage of collecting range and intensity information in a single frame, assuming that the ROICs are made to support this (making them true *flash imaging* systems). These systems usually operate at lower frame rates, with a high probability of detection on each frame (~ 30 Hz), requiring higher laser energy per pulse compared to GMAPDs. This makes most LMAPDs less compatible with efficient fiber lasers, which are peak-power limited.

Historically, LMAPDs have not been as sensitive as GMAPDs, requiring additional total laser energy per sample. While LMAPDs have the advantage of being able to make a 3D point that can image on a single laser pulse return, a more sophisticated ROIC is required on which to capture and store this imaging data from each detector. Because of this sophisticated ROIC requirement, a larger pixel size is likely to be required compared to that of a GMAPD array camera.

Polarization-based Flash Lidar using Framing Cameras

For **polarization-based flash lidar imaging** using framing cameras, temporal (range) resolution is provided by a **Pockels cell**, as shown below.



Light enters the receiver. A single polarization of return light is isolated. Alternatively, twice as many cameras can be used to detect both polarizations. A ramp is placed on a Pockels cell to switch polarization as a function of time. Multiple ramps are shown in the figure to depict **unambiguous range**. Two standard framing cameras are used.

In any given detector, the ratio of the power in one camera to that in the other camera provides range information. The primary disadvantage is that a Pockels cell is needed to rotate the polarization, and Pockels cells traditionally require high voltage and have a narrow FOV. A steeper slope provides more accurate **range information** but also has a smaller unambiguous range.

A number of standard techniques can be used to expand the unambiguous range. One of these is **chirping** the length of the ramps. The primary advantage of chirping is that a pair of **standard framing cameras** can be used for **high range resolution**.

Laser Vibration Detection

The **Doppler shift** changes the frequency of the return laser signal based on the ratio

$$\Delta f = \frac{2V}{\lambda}$$

where *V* is velocity, and λ is wavelength.

Velocity (μm/sec)	Frequency (Hz)
1	1.29
10	12.9
100	129.03
1000	1290.32

The table shows Doppler frequency change due to **vibration** for a wavelength of 1.55 μm. As an object vibrates, it creates velocities toward and away from the sensors. Vibration creates a FM beat signal between

the LO and return signal, with the modulation frequency equal to the vibration frequency, and the magnitude of the frequency shifts equal to the varying Doppler shifts. For high Doppler sensitivity, detection is usually performed using **temporal heterodyning**, where the return signal is combined on the detector with a LO to create a frequency-downshifted signal.

A **laser vibrometer** can characterize many types of vibration. For example, it can identify the type of engine in a vehicle based on the engine’s vibration. It can easily differentiate between a turbine engine and a piston engine, or determine the number of cylinders in a piston engine. Almost any type of vibration can be identified. The **vibrational image** of an internal combustion engine can be used to identify combustion pressure pulses and inertial acceleration of the pistons and drive trains. It can also help identify mechanical imbalances and misfires. Power through a transformer or liquid through a pipe can also be characterized.

Flash illumination for vibration detection has the potential to be a powerful technique because it allows simultaneous vibration measurement over an area. However, it requires a 2D array of photodetectors operating at very high frame rates. Such imaging cameras currently exist at low frame rates, in the range of a few hundred hertz.

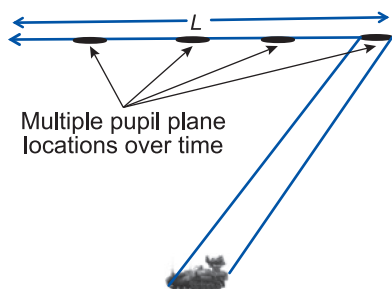
Synthetic Aperture Lidar

Synthetic aperture lidar (SAL) uses the motion of the lidar to develop an effective aperture that is larger than the real aperture. At a given location, the lidar emits a pulse and measures or estimates the returned field from the target. If the returned field can be measured at many different locations, those fields can be stitched together in the pupil plane to develop a larger pupil plane image, which can be Fourier transformed to obtain a high-resolution image.

It is not surprising that **synthetic aperture imagers** were developed first at microwave frequencies, where the wavelength might be 3 cm, such as with 10-GHz **synthetic aperture radar** (SAR). It is easier at longer wavelengths to align pupil plane images to a fraction of a wavelength. In addition, the real aperture angular resolution of lidar is very good, so the need for SAL was less in the past, whereas SAR has existed for decades. Generally, in SAR, the point of closest approach, called the **phase center**, is used as zero in the reference system. Because of the reference frame adjustment and, more fundamentally, because both the transmitter and receiver move, the resolution in the along-track direction for a SAR is given by

$$\Delta\theta = \frac{\lambda}{2L}$$

where L is the distance flown. This is twice the resolution of a real aperture with a diameter the same as L .

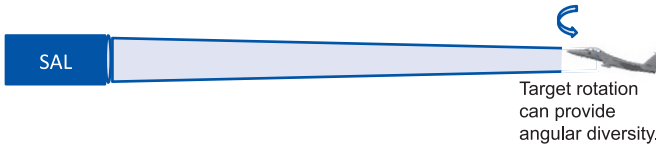


For a 1.5- μm SAL, the required distance flown will be 20,000 $\times$ shorter than at 10 GHz, but the motion compensation precision will be 20,000 $\times$ more precise. SAL should be able to provide long-range, high-resolution

imaging in angle/range space. Range resolution is $\Delta R = c/(2B)$. Holding the 2D resolution constant, SAL systems have power requirements that scale as $1/R^3$.

Inverse Synthetic Aperture Lidar

Inverse SAL (and **inverse SAR**) is the same as SAL (and SAR), except that the **angular diversity** is obtained by motion of the object being viewed instead of by motion of the sensor. The same basic physics is at play.



In SAL, the target stays oriented in the same direction, but the lidar moves, creating an **angular rotation** that allows for angular diversity. The sensor moves a distance L and is a distance R away from the target. The resolution is, again,

$$\Delta\vartheta = \frac{\lambda}{2L}$$

If the target is a distance R , the resolution obtained is $R \times \Delta\vartheta$. If, instead of the lidar moving, the target rotates,

$$\Delta\phi \approx \frac{L}{R}$$

Here, we have the same angular diversity and can obtain the same resolution. Historically, engineers have tested SARs by looking at targets placed on a rotating platform. The physics is the same as in SALs. The pupil plane field is captured multiple times with the target in multiple orientations. A reference frame adjustment is made, and a larger pupil plane image is created. This image is Fourier transformed to obtain the object plane image of the target.

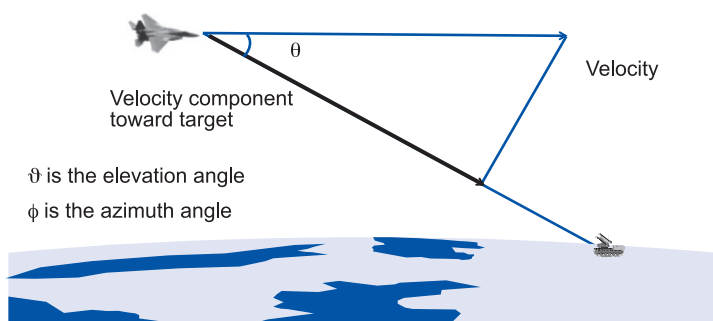
In SAL, the **reference frame adjustment** involves adjusting each pupil plane image with the illuminator maintained in the same place. In inverse SAL, the reference frame adjustment involves adjusting each pupil plane image with the illuminator maintained at the same angle with respect to the target.

Range Doppler Imaging Lidar

A microwave radar can make a 2D image using only one detector. A lidar can do this as well, but the FOV is much smaller because the diffraction-limited spot size is much smaller. Page 3 discussed range resolution being available from a single detector based on the speed of light and the time required for a laser pulse to travel to the target and return. Assuming that the lidar is closing on the target, the Doppler shift is, again,

$$\Delta f = \frac{2V}{\lambda}$$

This value of frequency shift applies only to an object going directly to or from the lidar.



The velocity toward the target is

$$V_T = V \cos(\theta) \cdot \cos(\phi)$$

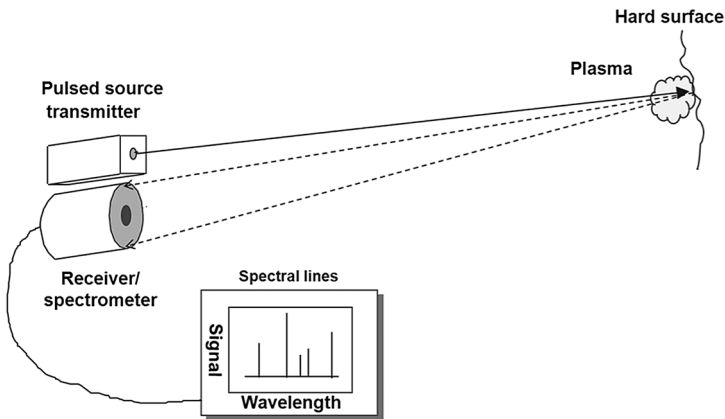
The change in velocity across a given detector can provide azimuth information without the need for multiple detectors. This technique is called **range Doppler imaging** because the image can be made using only range and Doppler information obtained from a **single detector**. This works very well for radar because the radar beam width is so large compared to lidar; however, for lidar, this technique will provide a very small image. In order to make the range Doppler image larger, a larger area must be illuminated, and a detector array must be used.

Laser-Induced Breakdown Spectroscopy

Laser-induced breakdown spectroscopy (LIBS) uses intense pulses from a laser to induce **optical breakdown** in a material, typically a solid, and then detects the **emission spectrum** from the breakdown. For nanosecond-duration pulses, the initial breakdown creates a plasma that then absorbs a large portion of the energy in the pulse. In the femtosecond-to-picosecond range, the shorter pulses are over before the plasma absorbs more of the pulse energy.

Recent advances in LIBS technology employ two pulses, spaced apart by 1 to 10 μs , where the second pulse acts to heat the plasma created by the first. The detection system then observes the emission created from the **irradiation** process, with **time gating** of the detector to optimize the desired emission spectrum signal. If the pulse or pulses are powerful enough, the emission contains components from the electronic transitions of individual, ionized atomic constituents of the irradiated material.

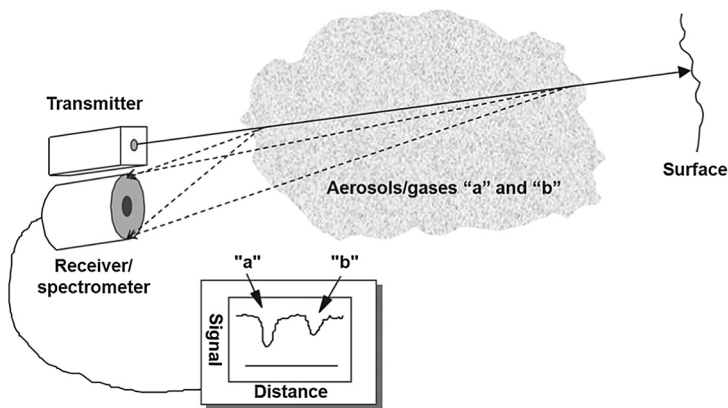
Through **spectroscopic analysis** of the **emission wavelengths**, typically in the 170- to 1100-nm range, one can determine the **atomic constituents** of the material, using the characteristic wavelengths of their emission lines. In addition, from the strength of the signals at each line, it is possible to determine their relative concentration, and from this, the chemical makeup of the irradiated material.



Laser-Induced Fluorescence Lidar

Laser-induced fluorescence (LIF) lidars detect material in the atmosphere or on the surface of a target of interest. LIF systems seek to create transitions by exciting material to higher electronic levels. In some cases, the material returns to the ground electronic level, in part by emitting light (fluorescing).

A LIF lidar system employs a receiver tuned to the wavelength of the fluorescence. Through **time analysis** of the fluorescence, it is possible to localize the detected material in the atmosphere, thus providing **range-resolved data**, and by employing a tunable receiver, or multiple receivers, one can detect multiple materials (“a” and “b”) as a function of range, as shown.

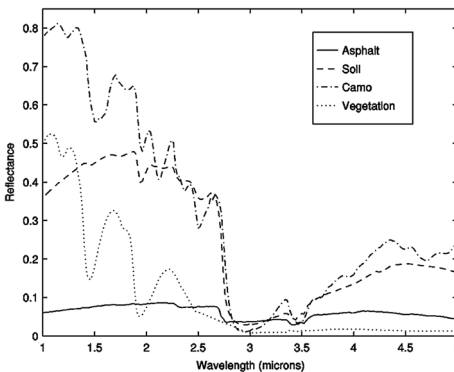


It is also possible to detect fluorescence from materials on a surface. Typical systems employ **pulsed ultraviolet sources** for the required excitation of the **electronic levels** as well as **range/species resolution**. The decay time of the returned signal, as well as the fluorescence spectrum itself, can provide some ability to distinguish different materials.

Active Multispectral Lidar

Color can be a useful discriminant. Color distinctions are based on the difference in reflectivity versus wavelength. **Active multispectral imaging** of targets has an advantage over passive imaging in that the illumination source can be controlled in active imaging. Even at night, near-IR wavelengths can be used, whereas, there is not enough signal to use passive multispectral imaging at these wavelengths at night. Using active imaging, one must also control both the spectrum and direction of the light source (as well as enable range resolution for obscurant and foliage penetration). An active multispectral sensor combines the benefits of conventional lidar and multispectral wavelengths in a band that has significant color variation in its reflectance.

Conventional lidars and some passive imaging sensors utilize the shape of a target for detection and/or identification purposes. Two different approaches have been used to deploy active multispectral sensors against **hard targets**. The first uses the laser wavelengths that are easy to generate, e.g., $1.064\text{ }\mu\text{m}$ and $\sim 1.5\text{ }\mu\text{m}$, and takes whatever recognition benefit can be gained. The second uses wavelengths determined to make the best active multispectral recognition probabilities and lasers appropriate to these discriminants. A class separability metric is formulated, and **optimal wavelengths** are selected.



The plot shows reflectivity versus wavelength for four materials. Clearly, customized wavelength selection is more effective than selecting a few readily available wavelengths to discriminate between materials.

Lidars Using Polarization as a Discriminant

Light is a transverse electromagnetic wave (i.e., the oscillations are perpendicular to the direction of propagation), similar in some respects to the waves that propagate down a taut string. The wave can oscillate horizontally, vertically, in a right- or left-handed spiral, or in some combination of the four. Mathematically, the four components of the polarization are represented by a set of quadruples known as the **Stokes vector**.

With active multispectral lidar, one controls the **polarization** and **direction of the illumination**, enabling, in principle, full **Mueller matrix characterization** of the target. Multiple measurements with filters rotated with respect to each other can identify the **degree of polarization** (DOP) of the light received. **Surface roughness** seems to be one of the parameters that influences the ratio of light reflected at different polarization states.

Passive DOP measurements are limited by the uncertainty in the initial polarization state of the ambient illumination and the angle of incidence of the light. In addition, if more than one material is present within the FOV of a single pixel (e.g., when leaves or camouflage are in front of a vehicle), the intensity and the DOP measured will be the result of the summation of all of the surfaces within the FOV. Because these proportions are not known, the **mixed-pixel effect** further adds to the uncertainty.

Active EO **polarimetric measurements** can eliminate this uncertainty by controlling the direction of illumination and the initial state of its polarization. In addition, signal generated by foliage or camouflage can be separated in range from the signal generated by whatever lies behind or in front of the object of interest. This minimizes the mixed-pixel effect unless the two surfaces are separated by less than the range resolution of the lidar.

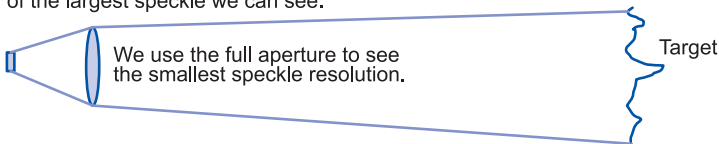
Speckle Imaging Lidar

When a beam of light from a laser transmitter illuminates a target, **phase irregularities** occur in the backscattered light due to surface roughness on the scale of a wavelength of light. **Interference** between the various contributions to the optical field produces a **speckle pattern** of bright and dark intensity regions in the receiver. This surprised early laser researchers because we do not see speckle in our everyday life. Our eye sees broadband light, averaging the interfering returns so we do not see light and dark areas.

In many optical applications, speckle is considered a nuisance (e.g., it degrades target images for narrowband lidars). However, researchers since the 1960s have known that the speckle pattern carries information about the physical properties of a target.

The interference we call speckle only comes from the region of an illuminated target, so the finest speckle return seen in angle space is based on either the size of the target or the size of the illumination area, whichever is smaller.

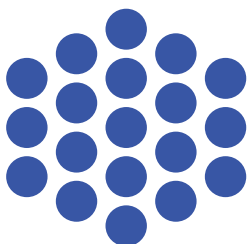
Our sensor resolution limits the size of the largest speckle we can see.



The full lidar aperture diameter on receive then limits the finest speckle angular resolution we can see, although the DAS can limit the finest observable angular resolution if the lidar is not diffraction limited. The achievable resolution in a **speckle imaging lidar** system is limited by the spatial extent of the pupil plane sampling. This can be thought of as the diffraction limit of the speckle imager.

Phased Array of Phased-Array Imaging Lidar

Optical phased arrays (OPAs) are devices that steer a laser beam and have no moving parts. They are space fed, meaning that the laser beam is generated, and the phase is imposed across that beam to steer light, as compared to phased arrays, which are made up of individual transmitters/receivers that are each only a fraction of a wavelength across. Nonmechanical steering approaches such as OPAs can be used in an array of subapertures.



Receive-only array
of subapertures

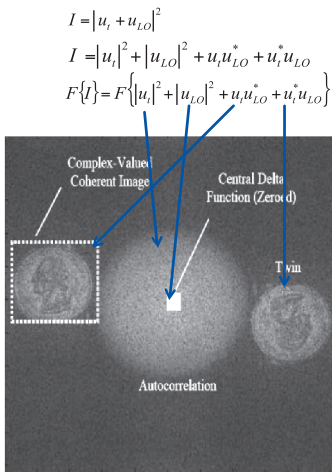
Receive-only arrays use a single transmitter, unlike **transmit and receive arrays**, which are referred to as **multiple-input, multiple-output** (MIMO) lidars. A MIMO array can have many transmit subapertures of one size and set distribution, as well as many receive subapertures. The receive subapertures can have diameters and distributions that differ from those of the transmit arrays.

The author groups all of these **multiple-aperture lidars** into a category called **phased array of phased-array (PAPA) imaging lidars**. This terminology is used in portions of both the lidar and directed-energy community, so the term is introduced here.

- A major benefit of PAPA arrays is that the focal length of each subaperture is directly related by the F-number to the subaperture diameter. This focal length is much smaller than that of a single, monolithic aperture with the same F-number. As a result, the size and weight of a PAPA array is much smaller than the size and weight of a monolithic aperture.
- The disadvantage is that the multiple subapertures now need to be phased up. Phasing up the subapertures would generally require a narrowband laser and some sophisticated algorithms.

Multiple Subapertures on Receive for Lidar

Using a single transmit subaperture and many receive subapertures has the benefit of a much shorter focal length than using a monolithic aperture, resulting in less volume and weight. This arrangement has the disadvantage of requiring phasing between the receive subapertures. To phase up the subapertures, one needs to either measure or estimate the pupil plane field in each subaperture. These measurements can be in the image or pupil plane. To form the high-resolution image, a pupil plane image is needed for the full array. That pupil plane image is then Fourier transformed to obtain the high-resolution image.



Either **spatial heterodyne** or **temporal heterodyne imaging** can be used to measure the field in each subaperture. Spatial heterodyne imaging is currently easier because it can use low-bandwidth framing cameras available in larger formats. It can measure the field across a given subaperture, but estimating the field of the entire array of pupil plane images requires optimization of a sharpness metric to estimate the phase differences between

subapertures. This is sometimes called a **hill climbing algorithm**, where a metric is chosen and optimized with multiple attempts to sharpen the metric, similar to climbing a hill.

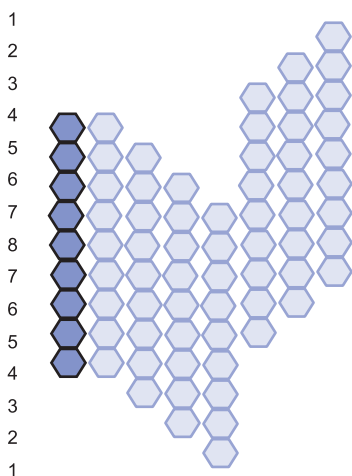
Stochastic parallel gradient descent (SPGD) is one of the more popular optimization techniques. Motion from one snap to the next is one of the issues to overcome in phasing subapertures. That along with requiring multiple speckle realizations makes phasing these subapertures challenging. With high-bandwidth, temporal heterodyne, pupil plane images, it would be possible to phase the subapertures without a hill climbing algorithm.

Multiple-Input, Multiple-Output Lidar

Using multiple transmit apertures and multiple receive apertures allows design flexibility not present in monolithic apertures. Instead of using motion to synthesize a larger effective aperture, as in a SAL, **MIMO lidars** use multiple physical **subapertures** to create a larger effective aperture.

An array of receive-only subapertures can synthesize an **effective aperture** as large as the receive array if the field across the array is known. With an array of transmit and receive subapertures, even more flexibility is obtained, provided it is possible on receive to identify from which transmitter each photon initially originated.

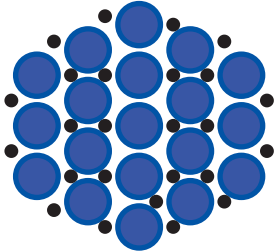
One effect of multiple transmit and receive subapertures is increased **angular resolution**, similar to the angular resolution increase of SAL. If transmission occurs from a subaperture in the middle of an array, the receive aperture array is effectively in its normal location. If the transmit beam is moved up one subaperture, the effect is as if the receive aperture were moved down one subaperture.



In the figure, the dark color shows actual locations of the arrays. The lighter color shows perceived locations on receive due to changing the transmit location. The number on the left is the number of received subapertures perceived to be at that location if all of the subapertures are used as transmit apertures. The fact that the number changes results in an **effectively apodized aperture array**.

Methods of Phasing MIMO Lidars

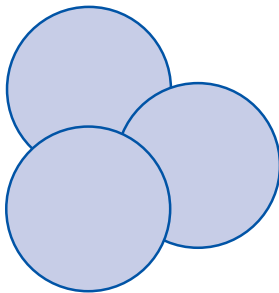
Phasing arrays with multiple transmitters and multiple receivers can actually be easier to align in phase than phasing arrays with only multiple receivers. Ideally, multiple transmit subapertures will emit a unique code so that on receive it is possible to determine from which transmit subaperture a photon originated.



Array of receive (larger) and transmit subapertures

The figure shows small transmit subapertures and larger receive subapertures, consistent with a flash imaging lidar implementation. In order to phase up the pupil plane image across all of the receive subapertures, we need to transform the receive images, adjusting each received array as though a single transmitter had been

used at one location. Each transmitter location needs to be transformed as though it originated from the reference transmitter location. For each transmitter except the reference transmitter, the reference frame on receive is adjusted as though the pupil plane were received in a different location. This provides multiple images of the same location, which can be used to closed-form calculate the phase difference between subapertures, even without a direct phase measurement.



Overlapping pupil plane images

With a receive subaperture and three transmitters in different locations, we can measure a pupil plane field that has overlaps. In those overlap regions, the difference in phase can be seen, allowing a closed-form solution to the phase differences between subapertures. The hill climbing approach is also an option here, but now more tools

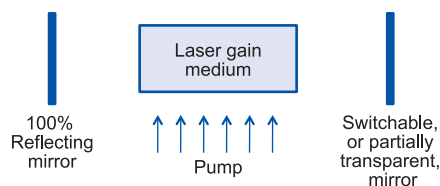
are available to **phase up the multiple subapertures**.

Lidar Sources and Modulations

In the 1970s, laser designators using **Nd:YAG lasers** came into use. A laser designator is simply a form of a **bistatic lidar**, where bistatic means that the lidar transmitter and receiver are in different locations. In the 1980s, the most

A Basic Laser

- needs a gain medium
- needs some method of pumping the medium
- need a resonator to obtain stimulated emission



popular laser for lidars was the **CO₂ laser**, usually operating at a wavelength of 10.6 μm , although a few CO₂ lidars operated on isotope wavelengths using carbon 13, or oxygen 18 instead of carbon 12 and oxygen 16.

These isotope CO₂ lidars avoided the 10.6- μm CO₂ absorption line in the atmosphere.

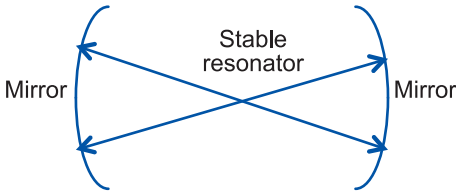
Currently, the most popular lasers for use with lidar are **solid state (SS) lasers** pumped by diode lasers. Often, pulsed lidars use a SS material that has a long upper-state lifetime so it is possible to Q switch the laser to create a nanosecond-class series of pulses. This requires fewer laser diodes and a lower cost to generate the same energy per pulse. High range resolution in a lidar requires short laser pulses or large bandwidth modulation. SS lasers can be Nd:YAG at 1.064 μm or various SS lasers operating at 1.5 μm to be more eye safe.

Very recently, **fiber lasers** have become popular for use with lidars. Fiber lasers can have more than 30% efficiency. They are restricted to lower peak powers because of the small cross section of the fibers. Fiber lasers in lidar generally operate at >10 kHz.

Diode lasers can also be used with lidar and are the most efficient form of laser. Fiber and SS lasers are usually pumped using diodes. Diode lasers do not always have narrow linewidths or a narrow spatial beam. Fiber or SS lasers can be thought of as **coherence converters**, taking lower-spatial- and temporal-coherence diode light and making a higher-quality laser.

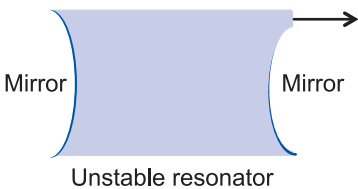
Laser Resonators

A resonator is required for a light beam to pass through a gain medium many times to create the amplified stimulated emission of a **laser resonator**. If positive amplification occurs each time the light is reflected back and forth in the resonator, stimulated emission develops, resulting in a laser. A simple stable resonator is shown. **Stable resonators** do not have



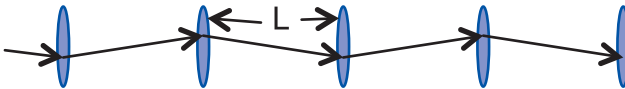
an output unless additional elements are inserted. For example, a polarizing element can be inserted, and a Pockels cell can be used to

switch polarization to create an output. In this case, the signal continues to increase until the output path is created and the energy is dumped from the cavity; this is **Q switching**. Some high-power lasers use **unstable resonators**. In this case, some loss on every path occurs, as some of the laser energy misses one of the end mirrors.



A resonator can be considered as a series of lenses in a row. The figure below shows a series of thin lenses separated by straight lines of length L , the length of the cavity. A full period of the

resonator is a half-lens followed by free space, followed by a half-lens, and then free space again.



This series of lenses can be modeled using matrices:

A thin lens can be modeled as

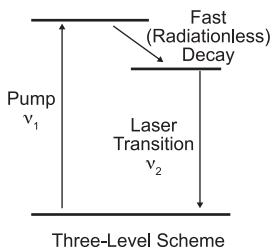
$$\begin{bmatrix} 1 & 0 \\ -\frac{1}{f} & 1 \end{bmatrix}$$

A straight section can be modeled as

$$\begin{bmatrix} 1 & L \\ 0 & 1 \end{bmatrix}$$

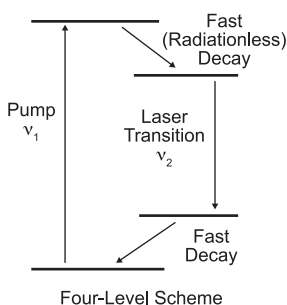
Three-Level and Four-Level Lasers

Population inversion is needed to achieve stimulated emission. Energy must be provided to pump a molecule from a ground state to an upper state.



In a **three-level laser**, a molecule is pumped from the ground state to an upper state. The upper state quickly decays by some nonradiative method to an intermediate state that is significantly higher in energy than the ground state.

Spontaneous and **stimulated emission** can occur between this upper state and the ground state. Ideally, only a small energy difference occurs between the initial upper state and the longer-lifetime upper state. The **quantum defect** is the amount of energy lost in radiationless decay; this loss needs to be minimized. The main issue with a three-level laser system is the high population in the ground state that occurs unless the laser is pumped very hard. Stimulated emission will be obtained only when the upper laser state has a higher population than the lower state. This means that the ground state must be depopulated, or absorption will result instead of stimulated emission.



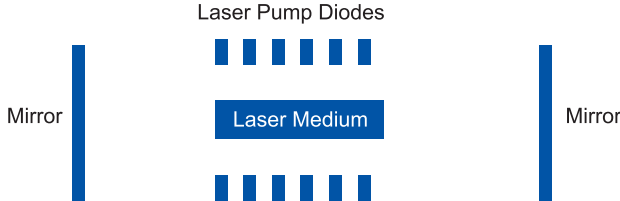
A **four-level laser** solves this problem by lasing between an upper state and a lower state that is slightly above the ground state. This increases the quantum defect loss; however, with fast decay into the ground state from the somewhat higher state, the population remains low. Four-level lasers do not require pumping as hard as three-level lasers.

Some lasers are called **quasi-four-level lasers** because their two lower levels are very close in energy; these lasers need to be cooled or else thermal agitation moves molecules from the ground state to this somewhat higher state.

Bulk Solid State Lasers for Lidar

Nd:YAG designators were some of the first lidars, albeit bistatic lidars. **Nd:YAG lasers** usually operate at **1.064 μm** . The eye cannot see 1.064- μm light, but it can focus it on the retina. The damage threshold is more than $10,000\times$ lower than light at 1.5 μm or longer, which cannot penetrate the eye to focus on the retina. As a result, a number of bulk SS lasers are used for lidar in the **eyesafe region** of 1.5 μm or longer. **Er:YAG** is a good example of an eyesafe laser around 1.5 μm . **Tm:YAG**, or **HoTm:YAG** are special cases of longer-wavelength lasers because a pump diode near 0.8 μm can generate two laser photons, decreasing the quantum deficit. Both Tm:YAG and Ho,Tm:YAG lasers operate near 2 μm .

Initially, bulk SS lasers were very inefficient and were flashlamp pumped. Over time, they became diode pumped, and the efficiency increased. Currently, the best efficiencies are about 20% wallplug efficiency. Various diode pump schemes can be used; one is **side pumping**:

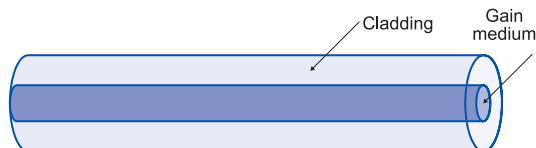


Early lidars using **bulk SS lasers** operated at low repetition rates of 30 or 60 Hz. These have worked well as framing cameras. Once flash imaging was developed, the energy per pulse went to hundreds of millijoules per pulse in a nanosecond pulse width. Long upper-state lifetimes of a SS laser are important for high-efficiency operation at low repetition rates because diode laser pumps are cw or quasi-cw devices. Nd:YAG has an upper-state lifetime of 220 μsec , so for any repetition rate above ~ 5000 Hz, it can use all of the energy from a cw pump. At 30 or 60 Hz, the laser pumps are turned on and off.

Quasi-cw operation means that the maximum power out of a diode can be increased by a factor of four over full-cw operation due to a cool-down period between pulses.

Fiber Lasers for Lidar

A typical **fiber laser** is cylindrical with a **gain region** surrounded by **cladding**. A **preform** is used, then the fiber is stretched in a drawing tower to make it longer but smaller in the cross directions.



A **diode** is typically placed at the end of the fiber, with a lens to focus light into the fiber and the cladding, although other pumping schemes can be used. Light often ricochets around both the cladding and gain medium, only pumping when it passes through the gain medium. A single-mode fiber has a small cross section, limiting the peak power available in fiber lasers. Recently, **photonic crystal fibers** have been developed; these allow higher peak power while maintaining a single mode. Average power is usually not a significant problem because the length allows for adequate heat dissipation.

Because of the **peak-power limitation**, fiber lasers are primarily used for higher-repetition-rate and cw lidars. The most prevalent peak-power issue is damage to the fiber tip; however, once light is in a fiber, nonlinear effects such as Brillouin and Raman scattering can be significant. These are not actually damage, but a distortion of the signal.

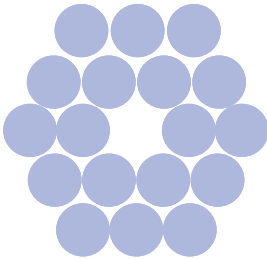
Fiber lasers can have 30% efficiency and are becoming increasingly prevalent in lidar not only because of this efficiency but because it is easier to wire a sensor together with fibers and couplers than it is to align free-space optics. Also, telecommunications uses fiber lasers, so many inexpensive and reliable fiber lasers and couplers are available, especially near 1.5 μm . GMAPD-based lidars use fiber lasers with ≥ 10 -kHz repetition rates. Materials with energy storage can be used in fiber format, so fiber lasers can be Q switched. Fiber lasers generally have very high beam quality because of their long, thin geometry.

Fiber lasers are ideal for GMAPD-based direct detection lidars and coherent lidars.

Higher-Peak-Power Waveguide Lasers for Lidar

The most stressing case for **lidar peak power** is a low-repetition-rate pulsed lidar for long-range flash lidar that illuminates a large area at one time. If you have 60- or 100-Hz operation and you want high range resolution, a lidar will have hundreds of millijoules per pulse in a nanosecond pulse. For long-range operation, it could be joules per pulse in 1 nsec—very high peak power—and the lidar usually needs at least moderate beam quality. For flash illumination, the laser beam can be spread out, mitigating required beam spatial quality.

Two approaches allow using higher-peak-power waveguide lasers for lidar: **photonic crystal fiber lasers** and **1D waveguide lasers**. Both approaches can have a single mode while increasing the cross section of the laser gain medium.



Cross section of a photonic crystal fiber

Photonic crystal fibers have a pattern of holes that can control the laser modes. The **1D waveguide** is constrained in a single dimension.

However, it needs a resonator in the other dimension unless it is only used as a power amplifier. Both of these approaches allow the cross-sectional area of a waveguide to become larger while still constrain-

ing laser modes. If a well-controlled laser mode is not required, the waveguide size will have no upper restrictions, but the spatial quality of the laser beam will be lost.

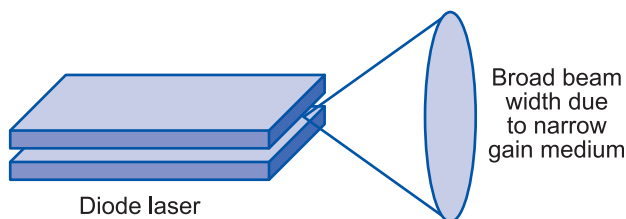
Lidars usually require **controlled laser modes**, although with flash illumination it is possible to reduce spatial coherence requirements on the laser source because a larger area is being illuminated. **Fiber tip damage** can be an issue, even with photonic crystal fibers, but the light should not need to be focused to as small an area compared to fiber lasers.

Multiple narrow-linewidth fiber lasers can be coherently combined as another method to obtain higher peak power; however, once combined, these lasers have the peak-power limitations discussed on page 65.

Diode Lasers for Lidar

Diode lasers are currently used in almost every lidar. In most cases, they are used as pumps for SS lasers. In some cases, they are used directly as the emitters for the lidar. Diode lasers have 50–80% efficiency and have the highest wallplug efficiency of any laser. (This is to be expected, since in most cases the other laser types start with a diode laser pump, so any loss makes other laser types less efficient.) Most diode lasers used for lidar are **interband lasers**, using transitions from one electronic band to another. Recently, **cascade diode lasers** have been developed for the MWIR and LWIR wavelengths.

Edge-emitting diode lasers have a thin gain layer in one dimension. This makes for high beam divergence in that dimension, meaning that fast optics are required for capturing the beam in the thin dimension. Also, if the gain region is allowed to be too large in the horizontal direction, the laser becomes multimode. Because of the one thin dimension, it is difficult to obtain high peak power from edge-emitting diode lasers.



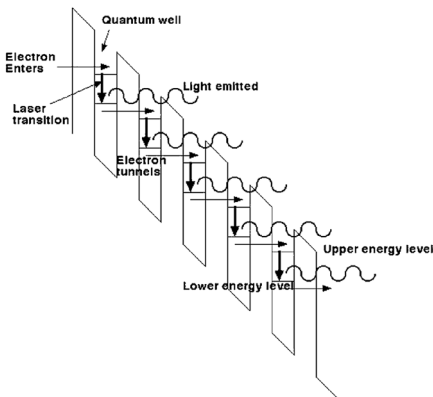
Vertical-cavity surface-emitting lasers (VCSELs) emit from the top surface rather than the edge, allowing better 2D coupling into a fiber, but also have size limitations due to mode control. Additionally, none of the diode laser materials exhibit long upper-state lifetime storage, so diode lasers are either cw, or quasi-cw, where, again, quasi-cw means approximately a factor of four higher peak power if the diode is allowed to cool down between pulses.

When using diode lasers directly for lidar, it is preferred to have **high-duty-cycle waveforms**. If the diode laser has a narrow linewidth, a diode laser with FM chirp or pseudo-random-coded waveforms can be used; however, it is difficult to make diode lasers with a narrow linewidth.

Quantum Cascade Lasers for Lidar

Lidar operating in the mid-IR band is now possible with **quantum cascade lasers** (QCLs). Power levels of QCLs are currently fairly low, but improving. Over time, QCLs will likely become one of the main IR laser sources for IR lidar applications. They do not have energy storage, so high-duty-cycle waveforms are required to make efficient lidars, since QCLs cannot be Q switched to obtain high-energy pulses.

QCLs rely on an **artificial gain medium** made possible by **quantum well structures**, through what is often referred to as **band structure engineering**. Common QCL devices are fabricated on InP and use GaInAs/AlInAs quantum wells, formed from 10- to 15-nm-thick layers. Because the energy separation of the states is a function of the structure, this bandgap can be adjusted to generate a wide range of wavelengths, covering approximately 3.5–30 μm and some of the terahertz region. For lidar applications, it is expected that the IR region will be of interest. It is possible that shorter IR wavelengths will become available with new materials.



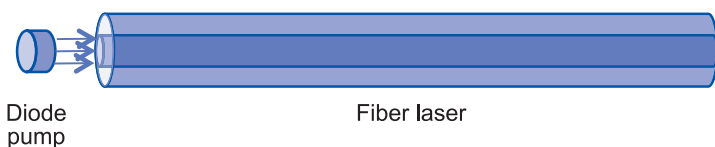
The term quantum cascade laser is used because of the “cascade” process involved, where one electron interacts with multiple quantum wells as it crosses the structure. The one electron creates multiple photons, corresponding to the number of quantum wells in the device, thereby

greatly improving efficiency. QCLs are able to operate in the MWIR and LWIR regions without the need for cryogenic cooling. QCLs can act as very effective seed lasers for nonlinear frequency conversion devices (optical parametric oscillators, etc.)

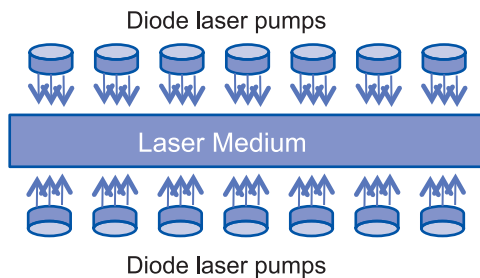
Laser Pumping Considerations

Laser diodes are electrically pumped, but most other lasers are pumped using light. Historically, **flashlamps** were used to pump lasers, but they are inefficient and unreliable. A flashlamp usually generates broadband light in a manner similar to a blackbody; however, most laser media capture only narrowband light, making flashlamps an inefficient way to pump a laser, even if the flashlamp itself is efficient.

Today, almost all lasers are pumped with laser diodes. Laser diodes can be both efficient and narrowband, so they can hit the spectral region that efficiently raises a laser medium to an elevated energy state, ready for stimulated emission. If the laser pump source is spectrally matched to the laser medium, the main issues for pumping efficiency are the efficiency of the pump source and the spatial overlap of the pump and the laser modes. **End pumping** can provide very high overlap efficiency between the pump and the laser mode but often cannot be scaled to high-power lasers. The figure shows a laser diode end pumping a fiber laser, with cladding around the gain media.



Side pumping allows many laser diodes to be used, converting incoherent laser diode energy into a spatially and temporally coherent laser power. Side pumping does not have as much spatial overlap as end pumping, so it is not as efficient. It can, however, pump much higher-power lasers.



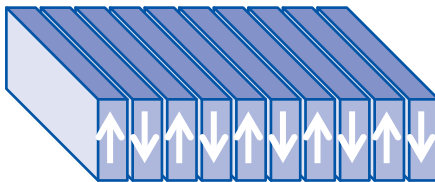
Nonlinear Devices to Change the Lidar Wavelength

It is often useful to have a lidar wavelength that is not the same as in existing lasers, or that can be generated more efficiently by using a high-efficiency laser at one wavelength and then changing the wavelength. The most common method to change wavelength is to use an **optical parametric oscillator** (OPO) or amplifier (OPA). One is a laser resonator and the other is an amplifier. Both take one high-energy photon and divide the energy between two new photons. The wavelength equation for this is

$$\frac{1}{\lambda_p} = \frac{1}{\lambda_s} + \frac{1}{\lambda_i}$$

where the subscript labels refer to the three waves: pump, signal, and idler, respectively. The output signal labeled *signal* (as opposed to *idler*) is chosen by the user.

To obtain the necessary **nonlinear effects**, high peak power or long gain length are needed in a material that has the right properties. The nonlinear conversion gain must be higher in an OPA than in an OPO because an OPA is a single-pass device. All three waves need to stay in phase so that interaction can occur. A decade or so ago, researchers began engineering materials that keep the waves in phase. The first to be developed was **periodically poled lithium niobate** (PPLN). In PPLN, when the waves begin to walk out of phase, the domains in the crystal are flipped so that they move back into phase. The figure shows a PPLN sample with



Crystal with periodically flipped (poled) domains

alternate domains flipped. PPLN layers tend to be thin, limiting the peak power out of an OPA. PPLN absorbs in the upper region of the MWIR, so periodic structures with broader trans-

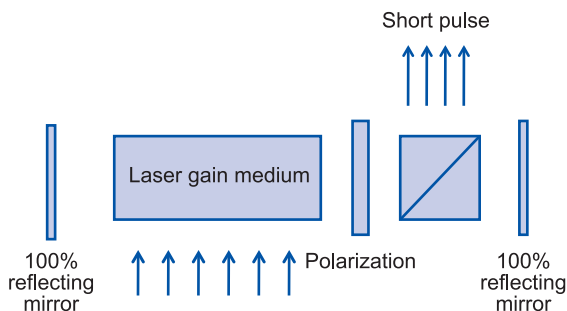
mission bands have been pursued. In addition to OPOs and OPAs, other nonlinear frequency conversion processes such as **sum and difference frequency generation** convert the wavelength of light.

Q-Switched Lasers for Lidar

A laser with a long upper-state lifetime can build up energy in a cavity, with no release. Then, suddenly, the energy can be switched out in a single pulse. The pulse can be nanoseconds long but can contain substantial amounts of energy, resulting in high peak power. This is called **Q switching** because we are switching the Q of the cavity.

Q is proportional to the ratio of the energy stored in the resonator at any given time to the energy lost from the resonator in a single round-trip cycle.

Q switching is a very effective way to efficiently generate high-energy nanosecond pulses. One method of Q switching is shown in the figure.



A **Pockels cell** can be used to switch from essentially *nothing* leaving the cavity during a round trip to *almost all* of the energy being suddenly dumped out of the cavity. A Pockels cell can be switched quickly, so it is commonly used for active Q switching. Passive Q switching methods use a **saturable absorber**, which is a material whose transmission increases when the intensity of light exceeds some threshold.

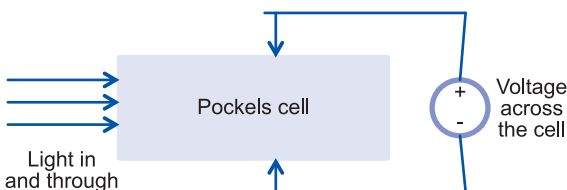
In order to Q switch a laser, the laser medium needs to store energy in the upper state. Lasers such as diode lasers or dye lasers cannot be Q switched. As a result, in general, nanosecond-pulse lasers are only used in lidars where the laser material allows significant energy storage in an upper state.

Pockels Cells

A **Pockels cell** is a crystal that changes the index of refraction of one polarization when a voltage is applied. Depending on the crystal, voltage can be applied in different directions. For the EO effect, birefringence is linearly proportional to the electric field. This is in contrast to the Kerr effect, which is usually smaller, and has a square relationship. At higher voltages, sometimes the Kerr effect can cause a larger **index of refraction change**, such as in KTN. Pockels cells can switch in a few nanoseconds, but they do require high voltage (often thousands of volts) and high current. They also can have a limited angle of acceptance. Engineers have looked at breaking up Pockels cells into multiple layers to reduce the required voltage per layer.

Many crystal types can be used as Pockels cells. The most common Pockels cell material is LiNbO_3 . The required voltage depends on the crystal used. Pockels cells are very useful for Q-switching lasers, and Q switching is very useful for lidars using a pulsed waveform. A Pockels cell will be used for Q switching, unless it is passive Q switching, which has less control. Also, Pockels cells can be used to switch the polarization of a laser signal. This can be useful for multiple lidar applications.

The **laser imaging and ranging** (LIMAR) approach to lidar ranging uses a Pockels cell to create a linear polarization ramp to obtain a range measurement without a high-bandwidth detector array. A Pockels cell can also be used for preventing damage to the optical receiver by isolating the receiver while a pulse is being transmitted.



Mode-Locked Lasers for Lidar

Mode-locked lasers provide periodic short pulses in regular series. Pulse separation τ_m is related to cavity length L by

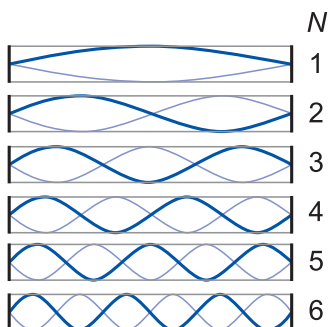
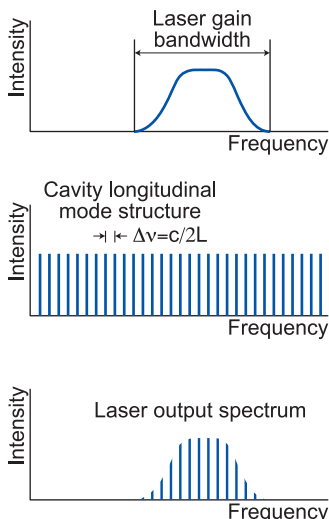
$$\tau_m = \frac{2L}{c}$$

where τ is the time needed for the light to make one round trip through the laser cavity. Mode-locked pulse width Δt is related to laser linewidth (depending on the specific pulse shape) by

$$\Delta t \approx \frac{0.44}{P\Delta f}$$

where P is the number of modes. The Fourier transform of the specific frequency line shape results in the pulse width. **Line broadening** can be homogeneous or inhomogeneous. If it is

homogeneous, the **frequency line shape** is Gaussian. If it is inhomogeneous, the frequency line shape is Lorentzian. Collisional or pressure broadening would be homogeneous, while the presence of dopants tends to promote inhomogeneous broadening.

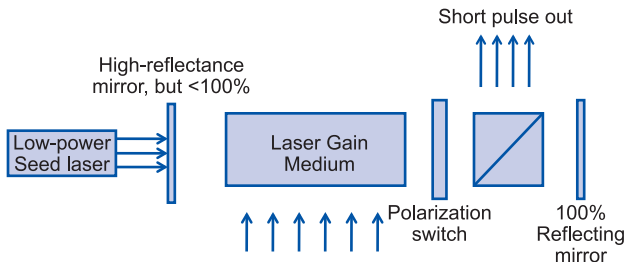


For N modes locked with a frequency separation $\Delta\nu$, the overall mode-locked bandwidth is $N\Delta\nu$. The wider this bandwidth the shorter the pulse duration from the laser. For example, the HeNe laser with 1.5-GHz spectral width has a pulse width around 300 psec, whereas the 128-THz-bandwidth **Ti:sapphire laser** has only a 3.4-fsec pulse width. Mode-locked lasers must meet the boundary conditions at the end of the laser cavity; only certain wavelengths are allowed.

Laser Seeding for Lidar

High-power lasers that are good candidates for a lidar often do not have the proper linewidth or beam quality. One way to improve **laser linewidth** and **beam quality** is to seed the laser. When **seeding a laser**, sufficient seed power is needed to cause the higher-power laser to follow the lead of the seed laser by locking to the lower-power signal instead of starting with random stimulated emission. In the absence of a seed, random fluctuations will build up in the laser cavity, starting with spontaneous emission. Stimulated emission can occur at any wavelength and in any mode consistent with the laser medium and the laser cavity.

Seeding a laser is a way to constrain the laser to oscillate with certain parameters. The process starts with the seed laser signal at a given level. The seed laser stimulates emission in the new laser, resulting in a signal that is locked to the incoming seed laser. Because the laser cavity is seeded rather than building up from random fluctuations, the laser signal has additional constraints beyond those of the cavity and the laser medium. The figure shows a seed laser inserted into a Q-switched laser.



Seeding can also be achieved with nonlinear processes. It is often performed with the signal line of an OPO to obtain a narrow-linewidth, seeded-laser output from the OPO. Seeding can encourage certain spatial modes, resulting in a narrow-linewidth laser that matches the linewidth of the seed laser. Seeding might be preferred over amplification for some laser media, e.g., when the laser gain in the medium is low.

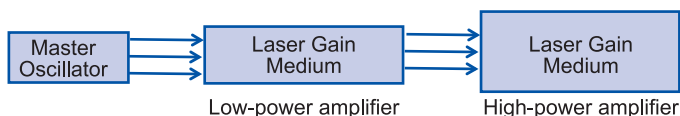
Laser Amplifiers for Lidar

Instead of seeding a laser to obtain the exact desired laser parameters, we can start with a **master oscillator** (MO) and amplify it. This configuration is often called a **master oscillator, power amplifier** (MOPA). The top figure shows a MO followed by an amplifier.



In this case, the amplifier is simply the **gain medium** with no resonator. The amount of amplification in a given stage varies depending on the length of the gain medium and its gain. Low-power amplification stages will likely have higher gain than high-power amplification stages. For high-power amplification stages, this process is more like power adding than linear amplification.

One issue with MOPAs is **isolation**. If the power amplifier is not isolated from the MO, feedback from the amplifier can distort the MO. Depending on the power of the MO, multiple **amplification stages** might be required to reach the power level desired, as shown.

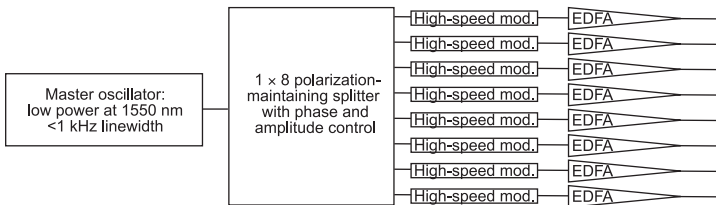


In a MOPA configuration, the efficiency of the MO and early-stage amplifiers has little effect on the **overall laser efficiency**. Even with, for instance, a factor of ten amplification in the high-power stage, only 10% of the power is consumed in the MO and early stages. Whether those early stages are 30% efficient or 3% efficient does not significantly influence the overall laser efficiency.

High-gain laser media are useful for amplifiers. For seeding lasers, the gain in the laser medium is not a significant issue because of the resonator. With high-gain media, significant signal gain can be achieved without a long amplifier stage.

Multiple Coherent Laser Transmitters

One lidar concept requires **multiple tagged laser transmitters** or a **transmitter array**. The user needs to be able to determine on receive from which transmitter the signal originated. The transmitters are coherent with respect to one another, meaning that a fixed-phase relationship exists between each one. A single MO signal with narrow linewidth (long coherence length) is used. The MO signal is then split into many different identical signals, and a unique tagging modulation is imposed on each of the signals. Each signal is then amplified [by erbium-doped fiber amplifiers (EDFAs) in the configuration below] to the level needed for lidar operation.



For coherent lidar, LOs are needed at each receive subaperture, requiring further splitting of the original MO signal. LO signals are significantly lower power than the transmitted signal, so they might not require amplification and can be frequency shifted.

In temporal heterodyne detection, it is common to frequency shift the LO signals, whereas in spatial heterodyne detection, the LO signals are generally not frequency shifted. MIMO techniques can be used with either spatial or temporal heterodyning. If **frequency shifting** is required, all of the LO signals can be shifted simultaneously. One would start with an MO signal, and split it into the LO signals and the transmitter signals. The LO line would be frequency shifted then further split into the required number of LO signals. The laser line would be further split into the number of laser signals, then tagging modulation would be imposed, and each signal could be amplified.

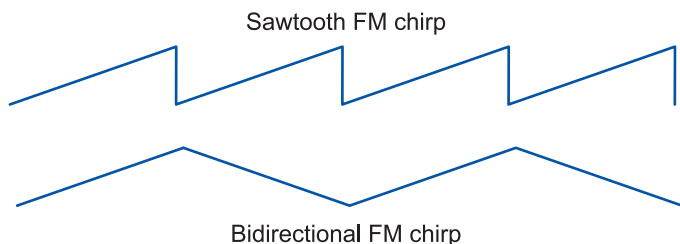
Laser Waveforms for Lidar

The simplest way to understand a **lidar waveform** is by looking at a single pulse. As a general rule, the speed of light can be considered as 0.3 m/nsec. A **laser pulse** travels to the target and returns, so a 1-nsec pulse provides 0.15-m range resolution. Range resolution is more precisely limited by the convolution of the pulse shape and the target, so pulse rise time can be a factor as well as pulse length. For a coherent lidar, velocity and range can be measured on a single pulse by using the **Doppler shift**:

$$\Delta V = \frac{\lambda \cdot \Delta f}{2}$$

where ΔV is the velocity resolution, and Δf is the change in frequency due to Doppler shift. The smallest measurable Δf is one over the time of the measurement; therefore, one difficulty with a short pulse that provides high range resolution is that its velocity cannot be accurately measured. A 1-nsec pulse would allow measurement of 1-GHz Δf , which, for a λ of 1.5 μm , would measure velocity down to 750 m/sec.

To measure high range resolution and high velocity resolution, a **pulse doublet** or other **high time-bandwidth product waveforms** can be used. Frequency modulation (FM) chirp and pseudo-random coded modulations are additional potential waveforms. With **FM chirp**, you can chirp in a single direction, or you can chirp both up and down in frequency, as shown.

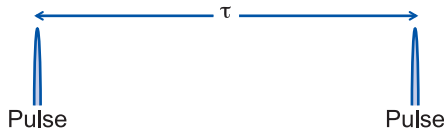


In **pseudo-random-coded modulation**, one randomly changes frequency or phase over the extent of the bandwidth.

The advantages of these waveforms are discussed in the remainder of this section, Lidar Sources and Modulation.

Polypulse Laser Waveforms

An ideal coherent lidar waveform allows for precise measurement of both range and velocity with a single waveform, called a **high time-bandwidth product waveform**.



Here is an example measurement that uses the high time-bandwidth product waveform:

Two 1-nsec pulses separated by 10 μ sec provide 0.15-m range resolution and 7.5-cm/sec velocity resolution for 1.5- μ m light. This has a **time-bandwidth product** of 10,000 because the velocity measurement capability has increased by a factor of 10,000 without changing the range resolution.

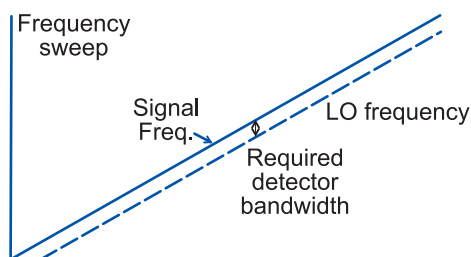
One way to obtain a high time-bandwidth product is to use multiple pulses in a waveform. The simplest **polypulse waveform** is two pulses, or a pulse doublet. A **pulse doublet** works as follows: the width of one pulse is used for range precision, and the width between the front of the first pulse and the back of the last pulse is used for velocity measurement precision. The phase measured in the first pulse must be related to the phase measured in the second pulse. If the phase has been randomized between pulses, the **doublet pulse waveform** will not allow precise velocity measurement. This means that the linewidth of the laser used must be less than one over the pulse separation.

More than two pulses can also be used if additional sampling is required to measure Doppler shift frequencies of interest. A frequency shift that is one over the separation of two pulses can be measured, but that is the only measurable frequency if the pulse is a delta function. To measure additional frequencies requires additional pulses between the two pulses.

Linear Frequency Modulation for Lidar

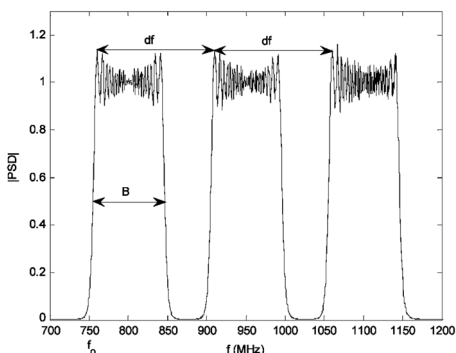
Range resolution is inversely proportional to the total signal bandwidth B . Normally, this would mean that a detector capable of responding to bandwidth is required; however, sometimes such a detector is either very expensive or impossible, depending on the range resolution indicated in the design.

One way to reduce the need for high-bandwidth detectors is to chirp the LO as well as the emitted signal. We use what



is termed **linear frequency modulation** (LFM) to obtain a smaller required detector bandwidth, as shown.

A second method of LFM uses multiple lasers simultaneously, each chirped a portion of the bandwidth required to obtain the desired range resolution. In this case, not only is a wide-bandwidth detector not needed, but a modulator that can generate the wide bandwidth is also not required. This method is sometimes called **stretch processing**, where the required receiver bandwidth is



determined by the **range depth** and **range uncertainty** (plus any LO offset floor to avoid $1/f$ noise).

This plot shows three chirped waveforms combined to be equivalent to a single larger chirp.

Pseudo-random-Coded Lidar

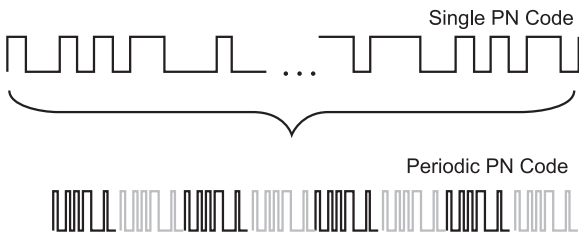
A **pseudo-random-coded waveform** has the same bandwidth requirement as previously provided for obtaining a certain range resolution:

$$\Delta R = \frac{c}{2B}$$

Unlike **LFM waveforms**, the required detector or modulator bandwidths in a pseudo-random-coded waveform cannot be reduced. However, the latter waveform can have more flexibility than an LFM waveform and can be easily generated. It is possible to modulate amplitude, frequency, or phase in a pseudo-random fashion; and, of course, frequency and phase modulation have a relationship. Also, if multiple transmitters are used, each transmitter can be assigned a unique, specific code.

Code division multiple access (CDMA) is an example of a pseudo-random-coded waveform and can be likened to people speaking the same language while other languages are perceived as noise and rejected. **Time division multiple access** (TDMA) is a second example and can be likened to people taking turns speaking. **Frequency division multiple access** (FDMA) is a third example and can be likened to people speaking at different pitches. **Pulse position (PPN) modulation** can have a single waveform or can have periodic codes. When periodic codes are used, one code length is called a **chip**.

PN in the figures stands for pseudo-random noise.

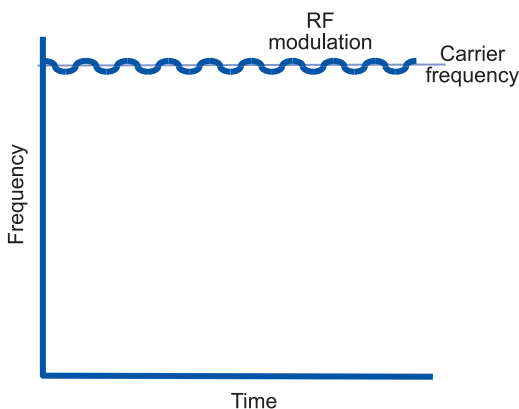


One issue with pseudo-random-coded waveforms is finding an “ideal” code. A code can be designed for optimum range resolution, optimum velocity resolution, or some combination of the two. Various algorithms address this issue.

RF Modulation of a Direct Detection Lidar

With the advent of laser communication over fibers, laser detectors with tens of gigahertz of bandwidth have been developed. We can therefore modulate a laser carrier to the X band and directly detect it. This is then a “simple” **direct detection lidar**, which does not use an optical LO, but it has some of the characteristics of a coherent lidar.

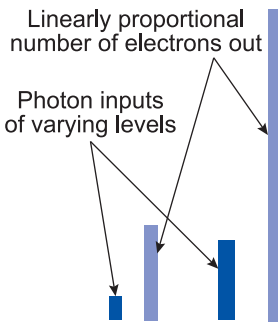
Once the signal is detected, it can be beat down to an intermediate frequency (IF) and processed the same way a microwave radar signal is processed. Doppler sensitivity will be based on the **RF modulation** frequency, not on the optical carrier. Range resolution will be based on the bandwidth of the RF modulation, unless another modulation is imposed in some fashion, such as using short pulses. Imposing an RF modulation on the optical carrier then using direct detection has the simplicity of the direct detection and can leverage radar IF and detection as well as processing techniques.



An optical carrier operates at 200 THz ($1.5\ \mu\text{m}$) or 300 THz ($1\ \mu\text{m}$), whereas RF modulation can be at the X band ($\sim 10\ \text{GHz}$) or the K_u band ($\sim 16\ \text{GHz}$). Lower RF frequencies can also be used. With high-bandwidth detectors becoming available, it might even be possible to use K_a -band ($\sim 35\ \text{GHz}$) modulation. The RF modulation imposed on an optical carrier can be FM modulated, AM modulated, or phase modulated.

Linear-Mode APD Arrays for Lidar

One way to reduce the effects of noise on a direct detection lidar is to employ gain either *before* the detector or *in* the detector. A fiber pre-amplifier can also be used. Historically, engineers have used image intensifier tubes. The **avalanche photodiode** (APD) is the SS equivalent of the photomultiplier tube in that both create many electrons from each incident photon. The number of electrons created per absorbed photon is the internal detector gain. The noise associated with the gain process is called **excess noise**. The gain directly improves the **SNR** in situations where system noise limits the performance.



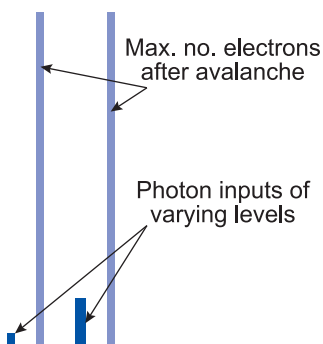
Currently, two major categories of APDs are in use: Geiger-mode avalanche photodiodes (GMAPDs; see page 83) and **linear-mode avalanche photodiodes** (LMAPDs).

LMAPDs can measure grayscale on a single pulse. The output is proportional to the reflected light, as shown in the figure.

An LMAPD can also provide a range profile from a single pulse, if range profile storage is built into the **ROIC** of the device. Building in this storage can make the ROIC physically larger and more complex than leaving it out. Responses from each range bin of interest must be captured and stored. Adding more range bins requires a more complex ROIC.

One of the challenges for LMAPDs is the ROIC development. Another issue is that the gain can be relatively low, so sensitivity will only be improved by a limited amount. Single-photon sensitivity for LMAPDs is underway but is developing more slowly than for GMAPDs. Typical LMAPD gain might be a factor of ten, but some state-of-the-art LMAPDs will have gain of 1000. High gain reduces the influence of noise on the detector.

Geiger-Mode APD Arrays for Lidar



The detector in **GMAPDs** is biased above the breakdown voltage so any photon hitting it will result in a large avalanche. If one photon, or many photons, hits the detector, the same large avalanche occurs, as shown in the figure.

Disadvantages of GMAPDs include:

- a **dead time** of perhaps 1 μsec after each triggered event, during which the detector will not detect any received photons
- a lack of ability to measure the signal intensity per pixel (or grayscale) on a single pulse
- possible cross talk between detectors
- possible dark current.

GMAPD flash imagers tend to run at high repetition rates because **low energy per pulse** is needed to obtain a response. Energy per pulse and probability of detection are kept low on a single pulse, then a number of pulses are integrated. This low-energy-per-pulse requirement allows for use of 1.06- μm lasers without eye hazard because of low single-pulse intensities.

An **effective grayscale** or an **effective range profile** can be developed by using **multiple pulses**. If multiple pulses return from a given location and the probability of detection is kept low, a higher reflectance area will have a higher probability of return, causing more events to trigger in conditions of high reflectance.

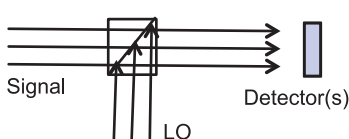
If pixels have mixed-range returns, the returns can be mapped as a function of range. If the probability of triggering is low for any event, (1) events will trigger at various ranges, and (2) there is a slight bias toward nearer ranges because of dead time after triggering.

A key advantage of Geiger-mode imaging lidar is the relative simplicity of the receiver and data acquisition electronics compared to wide-bandwidth, linear-mode receivers.

Receivers for Coherent Lidars

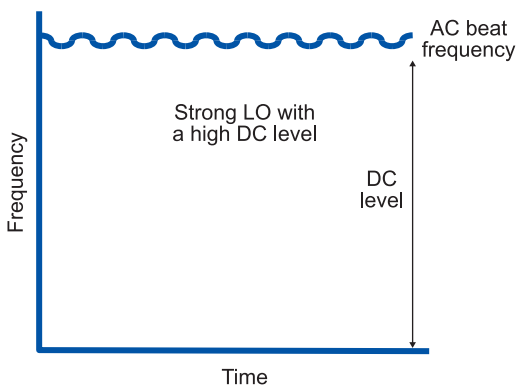
For a coherent lidar, the main method of reducing noise generally uses a high-power LO to beat with the return signal. If P_{LO} becomes large, the other terms in the denominator of the SNR equation are not relevant. Therefore, in general, gain (such as an APD) has not been used for **coherent detection**, although it can increase sensitivity.

Members of the radar community might need to become accustomed to the fact that in coherent lidars, the signal and the LO beat against each other by each illuminating the detector rather than beating against each other after detection; aside from this difference, coherent lidar detection is similar to radar detection.



Temporal heterodyne detection, where the LO and the signal are spatially aligned but usually offset in frequency, is shown in the figure.

In order for a high-power LO to reduce noise, one of the important receiver requirements is that it is AC coupled. If the receiver is DC coupled and the LO signal is large, the LO signal will kill the dynamic range of the receiver.



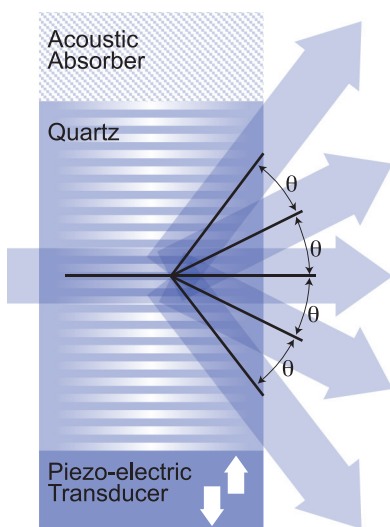
This figure shows a small AC signal on top of a large DC bias, similar to the large LO signal situation. From a noise-mitigation point of view, if a high-gain receiver is used, a strong

LO signal is not required. This can be important for reducing total required LO power in a large array.

Acousto-optic Frequency Shifting

Acousto-optic (AO) devices are commonly used to offset the LO from the emitted signal in temporal heterodyne lidars. They can beat the EM signal against a diffraction grating set up by the acoustic signal. This can be thought of as phonons beating against photons, and it can shift the frequency and associated wavelength of an optical signal.

For **temporal heterodyne detection**, an offset in frequency can help determine the direction of the object velocity; i.e., whether the Doppler shift is positive or negative. An **AO**



modulator (AOM) consists of a piezoelectric device that creates sound waves in a material such as glass or quartz, as shown. By vibrating the material with a pure sinusoid and tilting the AOM such that the light is reflected from the flat sound waves into the first diffraction order, up to 90% **deflection efficiency** can be achieved.

Thus, an AO device can deflect light, but it can also be used to shift the frequency of the light. Light is deflected according to

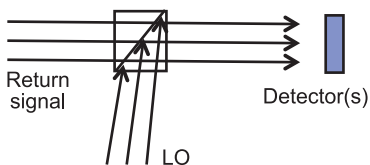
$$\sin \vartheta = \frac{m\lambda}{2\Lambda}$$

where λ is wavelength of the EM wave, Λ is the wavelength of the acoustic wave, and m is the order of the deflection, which, in a thin crystal, is ± 1.0 .

For heterodyne detection, the AOM frequency shifts the beam without deflecting it, or the beam deflection needs to be compensated for so that the LO signal and return signal both hit the target and overlap with high efficiency.

Long-Frame-Time Framing Detectors for Lidar

Two forms of lidar use **long-integration-time framing cameras**. One is **LIMAR**, a polarization-based lidar in which a Pockels cell provides timing information to



measure range. The second is **spatial heterodyne detection**, a form of digital holography, in which spatial phase is measured using an

angularly offset LO at the same frequency as the return signal, as shown in the figure.

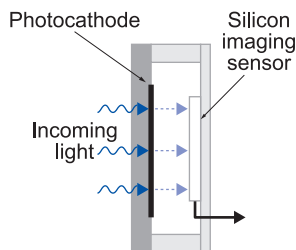
These applications use visible CCD or CMOS cameras, or they can use NIR, MWIR, or LWIR cameras, sometimes called FLIRs, or **forward-looking infrared cameras**. 2D cameras for lidar will probably be in the visible or the NIR region at about 1.0 or 1.5 μm . These cameras usually frame at 30 or 60 Hz. Unlike most lidar cameras, they have no high-speed circuitry but are very useful for these two types of lidar.

One of the major benefits of LIMAR and spatial heterodyne detection is their compatibility with widespread, affordable cameras with large format sizes. In the visible region, 10-megapixel cameras are available for hundreds of dollars. In the NIR region, cameras are more expensive, but still inexpensive, and with larger formats compared to high-bandwidth cameras. Megapixel NIR framing cameras are also available.

Because of the large format size for these cameras, along with the limited emphasis on single-photon sensitivity, it is likely that these lidars will be short range in order to have enough laser power to illuminate a large format image. The visible-range cameras are silicon and are very widespread. The NIR-range cameras are likely InGaAs, although they could be HgCdTe. The main disadvantage of HgCdTe is that it needs to be cooled.

Gated Framing Cameras for 2D Lidar Imaging

When using laser flash illumination, one aims to gate out noise except at a relatively small time period occurring approximately when the laser pulse returns. This is called **gated 2D imaging**. Range-gated imaging is performed primarily in the SWIR band with a 1.55- or 1.06- μm laser. This form of imaging started with an **electron-bombarded active-pixel sensor** (EBAPS[®]) based on a GaAs or InGaAs photocathode in proximity focus with a high-resolution, CMOS active-pixel sensor (APS) imager anode.



The SWIR version of this thermoelectrically cooled sensor has a spectral response that ranges from 950 to 1650 nm. Electrons emitted by the photocathode are directly injected in the electron-bombarded mode into the anode. Low noise gain is achieved in the CMOS anode by conversion of the high-energy photoelectron (resulting from 1- to 2-KeV bias applied between the photocathode and CMOS anode) to electron-hole pairs in the anode via the **electron-bombarded semiconductor** (EBS) gain process. The electrons are collected in the anode pixel and subsequently read out. The EBS gain process is inherently low noise, with an excess noise factor of less than 1.1.

A more recent gated-imaging system performs both gated 2D *active* and gated 2D *passive* functions in two spectral bands. This system uses an MWIR HgCdTe APD FPA with 640×480 -pixel format, 25- μm -pixel pitch. The MWIR APD FPA is fabricated using the HgCdTe technology described above. In the MWIR band, the **noise equivalent temperature difference** (NETD) at gain = 1 is ~ 18 mK (with $f/4$ optics and 16-ms integration), while in the SWIR band used for gated imaging, the **noise equivalent photons** (NEPh) value is ~ 2 photons for an integration time of 100 ns.

Lidar Image Stabilization

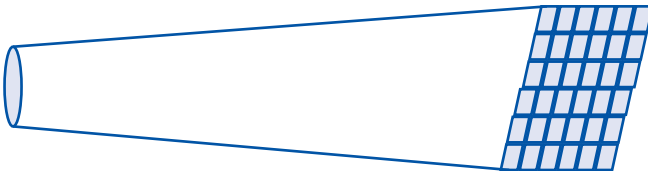
Stabilizing a camera when taking a photograph with high magnification prevents the image from being blurred; the same concept applies to the angle/angle portion of a lidar image. In a camera, less blur occurs with a faster shutter speed. A typical “shutter speed” for a lidar is nanoseconds, so image blur in lidar only occurs when pixels captured at different times are stitched together.

The main issue with **lidar image stabilization** is related to this stitching together of pixels captured at different times. In order to stitch together the returns from “images” taken at different times, one needs to know exactly where each image was taken to a fraction of that image size. If the snapshot at a given time is of a single pixel, one must stabilize to a fraction of the angular extent of one pixel. If the snapshot at a given time is of many pixels, one needs to stabilize only to a fraction of the angular extent of that snapshot.

The figure compares imaging one pixel at a time to imaging a 6×6 array of pixels. It shows that the array case does not need to be stabilized to the same level. If the full image is taken on one snapshot, no stabilization is required, except for maintaining the ability to be pointed at the target when the lidar image is taken.



Imaging a single pixel



Imaging an array defined by an array of detectors

One of the major benefits of **flash imaging with lidar** is the release in stabilization requirements. A camera takes the full image at one instant but has a longer shutter time, so objects can move during the detector integration time. Lidar imaging stops this motion.

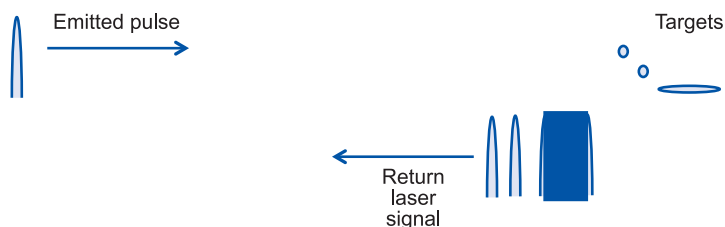
Range Resolution of Lidar

A lidar measures range by timing the travel time of a laser pulse to the target and return. As stated earlier,

$$\Delta R = \frac{c}{2B}$$

The factor of two is included because the laser pulse travels to the target and returns. When calculating **lidar range resolution**, it is helpful to assign the speed of light c the value of 0.3 m per nsec (1 ft per nsec). This means that a 1-nsec pulse will have about 0.15 m (6 in.) of resolution. Lidar pulses are often on the order of nanoseconds. If the lidar receiver has sufficient bandwidth to follow the return signal, the measured return signal for a pulsed lidar will be the convolution of the laser pulse and the target.

If the **lidar pulse** was a **delta function**, the range profile of the target will be captured exactly. The figure shows how the return signal is spread by the target range profile.



Not all lidars use a pulsed waveform. Some type of **time tag** is needed to determine the time needed for the laser pulse to travel to the target and return. This time tag can be provided by **frequency chirping** the lidar signal, if on receive the frequency chirp can be measured. It could also be provided by pseudo-random changes in frequency or in phase.

Range resolution is one over the signal bandwidth, regardless of how the bandwidth is changed. **LFM** and **pseudo-random-coded modulation** are used with **heterodyne detection**.

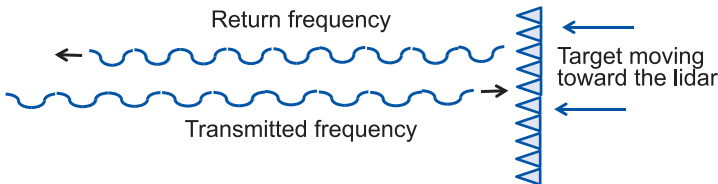
Velocity Resolution of Lidar

When a lidar pulse returns, its frequency will have been shifted based on the Doppler shift in the return light:

$$\Delta f = \frac{2V}{\lambda}$$

The faster the **velocity** to or away from the sensor the higher the frequency shift. The factor of 2 is included because of the two-way path of the light: to the object and returned from the object.

In astronomy, a **Doppler shift** occurs in the color of stars, but it is radiation traveling only one way, not two ways, so it would not have the factor of two. A star that is receding from the sensor is shifted to longer wavelengths, called a **red shift**. A star traveling toward the sensor is shifted to shorter wavelengths, called a **blue shift**.



Obviously, if people can see a red or blue shift, it is possible to measure the color and determine the Doppler shift. Direct detection lidars can measure Doppler shift with wavelength filters. However, if a very accurate measurement of the Doppler frequency shift is required, a coherent lidar is used. The coherent lidar measures the phase of the returned field using heterodyne detection.

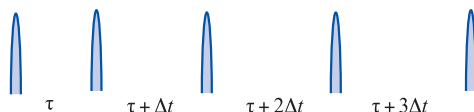
The frequency shift is the rate of change of a phase shift, so a Doppler shift also changes frequency. In order to see that frequency, the **Doppler shift measurement time** must equal one over the frequency shift. To obtain a better sample, the measurement time needs to be twice that length. This sets up the tradeoff between range and velocity resolution discussed on page 78. Because optical carrier frequencies are so high, the Doppler shift is large enough to measure even for very small velocities on the order of microns or tens of microns.

Unambiguous Range

If the range to an object is sufficiently long, two or more laser pulses can travel to and from the target simultaneously, possibly creating an ambiguity in range. For a fixed pulse repetition rate, the **unambiguous range** is given by

$$R_{\text{unambig}} = \frac{c\tau}{2}$$

where c is the speed of light, and τ is the time between pulses. The unambiguous range limits the distance an object can be while still knowing its range. When an object is farther than the ambiguity range, it is not known from which pulse the received return originated. One way to extend the ambiguity range is to change the time between pulses. Then, until the **pattern of changes** begins to repeat, it can still be determined exactly which laser pulse is being received. The pattern shows a small increment being added to each interval between pulses.



As long as the period continues to change, we can determine which pulse most recently traveled the entire distance from the sensor to target and back again, allowing for range measurement with no ambiguity. Range ambiguity will still occur but at a range that depends on when the pulse pattern repeats itself.

If the ambiguity range is 40 km with a fixed point spread function, and a minor change is made in the time between pulses for 10 consecutive pulses, the resulting unambiguous range would be approximately 400 km.

Unambiguous range is only relevant if the exact range of an object is critical. For tracking or imaging an object, it might not be necessary to resolve range ambiguity in order to determine absolute range.

Point Spread Function

The **point spread function** (PSF) is an angular spatial impulse function of the system. The PSF, along with the **detector angular subtense** [(DAS), which samples the angular PSF] allow the lidar to measure angle/angle resolution. The diffraction effects of the lidar receiver and the blurring effects of the atmosphere combine to produce the PSF.

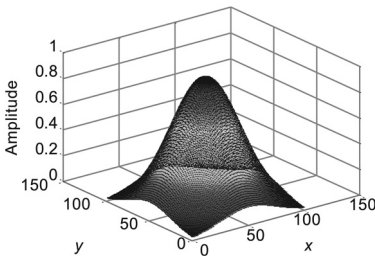
In most cases, the FPA is placed in the image plane to sample the PSF. Alternatively, the pupil plane field can be captured, in which case the size of the pupil plane sample, along with the atmosphere, will determine the PSF.

Perfect, nonaberrated imaging optics can be modelled as a **phase screen** with a **phase profile** of

$$t_{\text{lens}}(w_{\text{az}}, w_{\text{el}}) = e^{\frac{-j\pi(w_{\text{az}}^2 + w_{\text{el}}^2)}{\lambda f}}$$

where λ is wavelength, f is focal length, and the two w terms are the distance from the center of the lens, limited, of course, by the radius of the lens. This **phase function** brings the received pupil plane function to an image at a distance f from the lens. The smallest spot a lens can make depends on the diffraction limit of the lens. The **Airy disk** is

$$D_{\text{Airy}} = 2.44 \frac{\lambda f}{D}$$



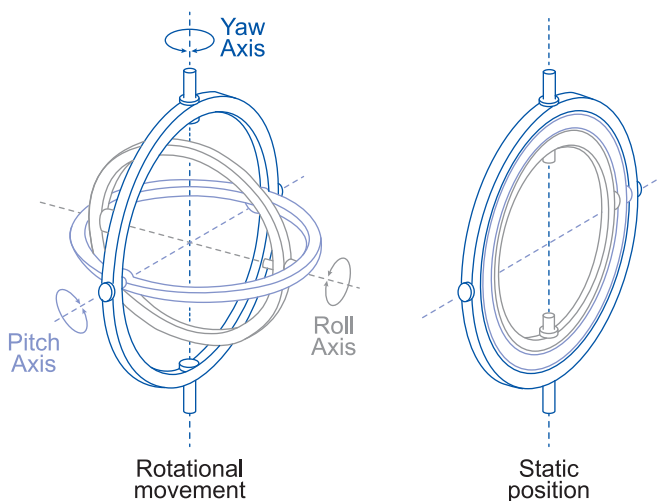
where D is the lens diameter. The Airy disk is the diameter out to the zeros of the diffraction-limited spot at the focus for a circular aperture. The focal length over D is the F-number of the lens. The DAS is generally set

at approximately the angular extent of D_{Airy}/f .

For a scanning lidar, the illuminated region can be smaller than the DAS. In this case, the angular resolution is limited by the illuminated spot size in angle.

Gimbals for Use with Lidar

Gimbals are mechanically complex and expensive but can steer very accurately over wide angles. Except for obscurations, a three-axis gimbal can point in any direction. Each axis can rotate, and using all three axes, a lidar aperture can be held in the inner ring and point in any direction.



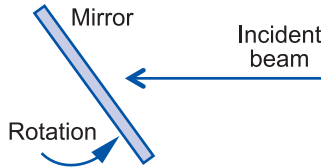
Obviously, the aperture inside the inner ring must be significantly smaller than the size of the outer ring. A hole that is $2.5\text{--}3\times$ the size of the aperture must be cut into the vehicle on which a gimbal is mounted.

One issue with gimbals is getting electrical or other connections into the inner gimbal. It is clear from the figure that it would be difficult to run a wire to the inner gimbal without the possibility of tangling the wire. **Slip rings** are sometimes used, where two conducting surfaces slide against each other to keep electrical contact while a ring rotates.

Most current military optical systems use **gimbal pointing**. Gimbals can usually be pointed and stabilized very accurately but tend to cost hundreds of thousands of dollars due to their complexity.

Fast-Steering Mirrors

A **fast-steering mirror** (FSM) is simply a flat mirror that can be moved very quickly over small angles. The beam is steered using the fact that the angle of incidence equals the angle of reflection as the mirror is rotated.



There is a tradeoff concerning the size of an FSM: The larger the mirror the more power will be needed to move it quickly. The size of FSMs is between those of a **MEMS mirror** and a full-sized **steering flat**. FSMs are rotating flat mirrors, but of different sizes. In order for FSMs to move quickly, they need to be relatively small, so they are placed where the optical beam is small. At the same time, the final steering angle θ_f is reduced by magnification M of the beam to the final aperture size:

$$\theta_f = \frac{\theta}{M}$$

where θ is the actual angle steered by the steering mirror. If the beam is very small where the FSM is placed in the system and the beam needs to be magnified, the steering angle is reduced by the magnification. In this case, a wide-FOV telescope is needed to magnify the beam to the final aperture size whenever an FSM is steering the beam in an area where the beam is small.

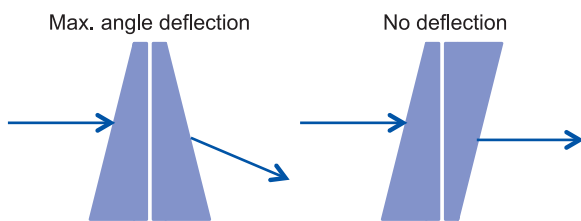
FSMs can move at ~ 1 -KHz bandwidth, so they can allow small-angle pointing fairly quickly. They have a **settle time** after moving quickly from one angular location to another. The settle time will likely be significantly longer than one over the FSM bandwidth.

Risley Prisms and Gratings

Risley prisms work by rotating two, or sometimes three, **steering prisms**. They can steer large angles using only **rotation**, a very simple motion. While the relative size of the aperture versus the size of the hole is an issue, it is possible to increase fill factor by placing the motor and much of the rotation mechanism in an area where the beam is smaller. Risley prisms also have a **zero-crossing** issue, meaning that if you track through zero, you need to rotate the wedges by 180 deg.

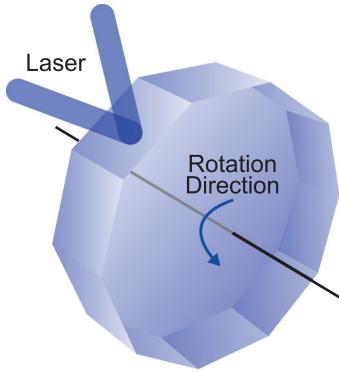
Risley gratings can solve this problem by switching the rotation from one handedness to the other. These gratings are also lighter than Risley prisms and can be moved more quickly, but they have dispersion. Dispersion will not be an issue unless the lidar is very broadband. For narrowband systems, Risley gratings are preferred over Risley prisms.

Risley prisms have the advantage of being a **conformal steering approach**, which means that they are flush with the surface they are mounted on, as opposed to gimbals, which can point over very wide angles but protrude into the wind stream. The figure shows two Risley prisms, positioned first for maximum steering, then for zero steering, but with some displacements of the beam.



Any time you steer using a mechanism that is flush with a flat surface, you have a cosine loss in power sent to a given angle, where the cosine is taken from a perpendicular direction. This cosine loss is due to the projected area of the aperture in the direction the beam is propagated.

Rotating Polygonal Mirrors



Rotating polygons have the following advantages:

- The only motion is simple rotation.
- Due to its many sides, rapid rotation is not needed to scan the angle rapidly.
- A scanning polygon has no flyback period; as soon as the beam crosses from one flat face to the next, the next scan begins.

Polygons scan in one dimension. Hexagons scan over 120 deg because the angle of incidence equals the angle of reflection. The optical deflection angle is twice the mechanical rotation angle. The more sides on the polygon the smaller the scan angle before jumping to the next scan.

A loss due to the size of the scanned beam does occur. As can be seen in the figure below, it will take time for the entire beam to traverse each corner. The size of the beam compared to the length of one side of the polygon provides a **scan efficiency measure**.



Polygon scanning has the following drawbacks:

- The polygon must be thicker than the clear beam and can become too large for small-angle scanning.
- Limiting the number of sides on the polygon scan causes efficiency to decrease.

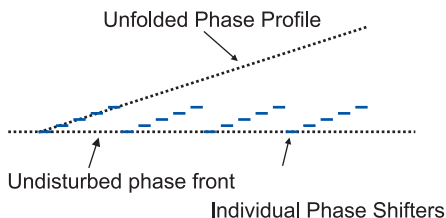
Modulo 2π Beam Steering

The index of refraction in a prism is larger than that of air, so light travels more slowly within the prism. The angle of light passing through a prism changes because the light moving through the thick portion of the prism will be delayed compared to light traveling through the thin portion.

Steering can be accomplished by changing the thickness of the prism. Light can be steered electronically by writing a prism. The problem is that it is difficult to create an **optical path difference** (OPD) as large as would be required to write a full prism of appreciable width. For example, a 10-cm-wide aperture steering to 30 deg would require >5 -cm OPD on the thick side of the prism.

For a narrow wavelength, we can take advantage of the fact that light is a sine wave. With sine waves, it does not matter if we have 0 , 2π , 4π , or $2n\pi$ phase shift; from a phase point of view, these are all the same. Therefore, moving across the width of the prism, one can subtract 2π of phase every time the phase reaches 2π , resulting in a **sawtooth phase profile**. The unfolded phase, which is called a **modulo 2π phase profile**, looks like the phase profile that would result from propagation through a full prism, and it steers light in the same manner.

The benefit of using a modulo 2π phase profile is that the required OPD can be small. The maximum required OPD approximately equals the wavelength of the light being used. The **modulo 2π beam steering** approach makes the beam-steering device very wavelength dependent (dispersive). The figure shows modulo 2π phase shifting to create beam steering.



Largest Steering Angle for an Optical Phased Array

The **largest angle that can be steered** to using the modulo 2π approach is determined by the size of the smallest individually addressable phase element. If the full array of phase shifters is illuminated with a Gaussian beam, any individual phase shifter will have an approximately uniform irradiance distribution across it. For a circular aperture, a uniformly illuminated full-width beam divergence at the **half-power point** is

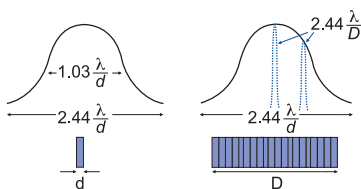
$$\theta \cong 1.03 \frac{\lambda}{d}$$

where θ is the beam divergence, λ is the wavelength of the EM radiation, and d is the width of the individual radiator.

If phase can be locked among many individual radiators, the beam becomes narrower in angle in proportion to the increase in the effective size of the radiator. If the full array is uniformly illuminated, $D = nd$ can be substituted into the equation, where n is the number of individual radiators assembled to make the large radiator, and it is assumed that the pitch of the radiator separations equals the width of the radiator (i.e., unity fill factor).

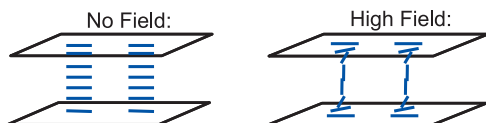
For **Gaussian illumination** of the full array, the effective size of the large aperture is reduced, and the beam divergence increases. The allowed amount of clipping of the Gaussian beam by the aperture array determines the degree to which the effective aperture size is reduced.

As shown in the figure, for **uniform illumination**, by adjusting the phasing among the individual elements, the narrow beam can be steered under the envelope of the larger beam resulting from an individual radiator.



Liquid Crystal Optical Phased Arrays

Birefringence of the LC is produced by the rotation of molecules under **applied voltage**. Alignment layers create forces that make one orientation preferred in the absence of an electric field. The figure shows a LC cell with voltage applied (left) and without voltage (right).



The speed of a **LC steering device** depends on how fast the LC molecules rotate from one orientation to another. The relaxation time τ_d for a LC cell to return to its no-voltage state, once the electric field is removed, is given by

$$\tau_d = \frac{t^2 \gamma}{k \pi^2}$$

where t is cell thickness, k is the effective elastic constant, and γ is viscosity. Cutting the cell thickness in half reduces **switching time** by a factor of four. Turn-off relaxation time is usually much slower than the time required to turn on the phase shift. Turn-on time is

$$\tau_r = \frac{\tau_d}{\left[\left(\frac{V}{V_t} \right)^2 - 1 \right]}$$

where V is applied voltage, and V_t is threshold voltage. Driving the turn on with higher voltage increases the speed by approximately the square of the ratio of increased voltage.

At high drive voltages, the turn-on time becomes much faster than the turn-off time. To speed up the turn-off response, engineers have developed **dual-frequency liquid crystals** (DFLCs), which drive the molecules both on and off. The dielectric properties of DFCLCs change sign as a function of the frequency of the applied AC voltage.

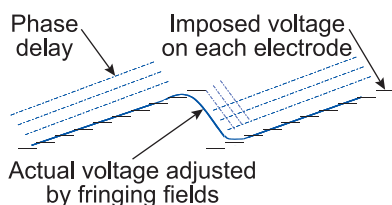
Another method to speed up LC beam steering involves mixing **polymers** into the LCs.

LC Fringing-Field Effect on Steering Efficiency

A **flyback region** is involved in performing a modulo 2π reset due to the fact that voltage across the LC cannot change sharply at the reset. During the flyback portion of the phase profile, the beam is deflected in the wrong direction:

$$\eta = \left(1 - \frac{\Lambda_F}{\Lambda}\right)^2$$

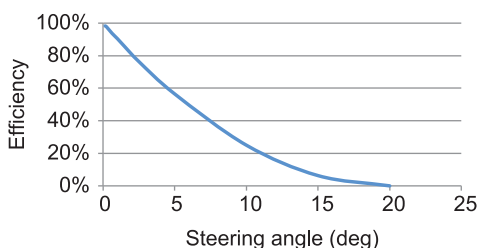
where η is efficiency, Λ_F is the width of the flyback region, and Λ is the width between resets.



Fringing fields expand outward to each side of the small electrode. As a general rule, the narrowest width of a voltage region above an electrode is approximately

the thickness of the LC layer between the electrode and the ground plane.

Because the thickness of the **liquid crystal layer** is often larger than the spacing between the electrodes, fringing fields can have a significant effect. The cell must be approximately thick enough to obtain one wavelength (2π phase) of OPD. With a birefringence of 0.3, this means that the cell needs to be about $3.3\times$ one wavelength in thickness.



Steering efficiency versus angle drops off very fast. If high steering efficiency is needed, the **steering angles** used for

sawtooth phase profile LC continuous steering need to be very small angles. For one-quarter of a degree angular steering, only about 98% steering efficiency would be achieved. For a 1-deg steering angle, efficiency would be down to 90%.

Reduction in Steering Efficiency Due to Quantization

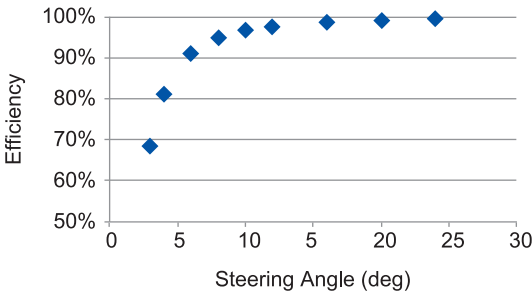
The second contribution to **steering efficiency** η is from the discrete nature of the **phase steps**:

$$\eta = \left(\frac{\sin(\pi/q)}{\pi/q}\right)^2$$

where q is the number of steps in the ramp. As can be seen in the table, very little quantization is lost when as few as 8 steps are reached, where 95% efficiency is achieved. 20 steps are required to reach 99% efficiency.

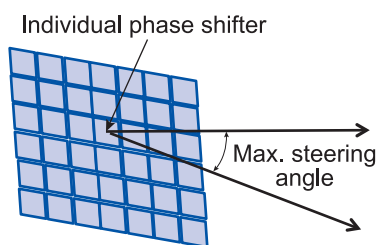
Number of steps	Efficiency
2	40.50%
4	81.10%
8	95.00%
12	97.70%
16	98.70%
20	99.20%
24	99.40%

In a LC beam-steering device, **quantization loss** is not as severe as either the equation or the table indicates. **Fringing fields** smooth out the steps, so a given discrete phase profile becomes blurred to appear more like a linear ramp, as shown in the plot of steering efficiency versus steering angle. Therefore, this loss is overestimated by the formula used for quantization loss.



Chip-Scale Optical Phased Arrays

The physics of OPAs is the same whether a chip or birefringent LCs is used to create the OPD. One chip is 64×64 , with a pitch of $9 \mu\text{m} \times 9 \mu\text{m}$, making the array $\sim 576 \times 576 \mu\text{m}$. The individual antenna radiators inside a pixel are $3.0 \mu\text{m}$ in length and $2.8 \mu\text{m}$ in width. A second OPA chip uses random element pitch to reduce sidelobes but is also a small chip. When phased-array individual elements are larger than a wavelength, **spurious sidelobes** result.



Another way to consider this is as follows: The largest full width, half maximum angle ϑ that can be steered to without significant sidelobes is limited by the size of the individual phase shifters according to

$$\vartheta = \frac{1.03\lambda}{d}$$

where λ is the wavelength, and d is the size of the individual phase shifter. Because we steer in only one direction and we want higher efficiency than 50% (half max), we need to restrict our steering angle to perhaps one-quarter of that value.

For example, if $3 \mu\text{m}$ is the size of the radiator, and $1.5 \mu\text{m}$ is the wavelength, we can steer to about $1/8^{\text{th}}$ radian, or about 7 deg. Let us assume that we would like to steer a 10-cm beam. This implies a magnification of about a factor of 174, reducing the 7-deg steering angle to 0.04 deg. However, over time, the steering arrays will become larger, and the individual steering elements will become smaller. Another issue is that these chips only transmit; they do not receive.

Nanofabricated phased arrays could soon become an interesting option for steering to small angles, using magnification after steering. The technology to make nanosized devices is rapidly improving.

MEMS Beam Steering for Lidar

Mirrors are dispersion-free and polarization-invariant components used in lidar. However, large single-aperture mirrors are bulky, slow, and consume high power. MEMS mirrors are much faster than non-MEMS mirrors because they have low mass and inertia. In order to form a phased array, 2π phase shift is needed between the MEMS mirrors.



Sawtooth phase profile from an array of micromirrors

A 2π phase shift can be provided either by positioning each micromirror or by adding a phase-shift layer before each micromirror. The phase layer needs only to be a π phase layer to provide 0-to- 2π phase shift because **MEMS-mirror-based steering** needs to operate in a reflective mode. The need to make the steering elements reflective can increase the size of a steering system. No flyback region occurs with MEMS mirrors.

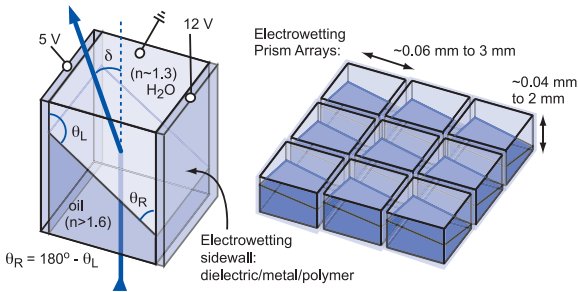
Due to size of the MEMS mirrors, the best method of steering is by **tilting** the micromirrors. No digitization loss occurs because the ramp is a single MEMS mirror. If either a piston motion or a separate phase layer can be used for phasing between phase ramps, this can be a very high-efficiency steering method. **Fill factor** will need to be addressed to keep high throughput efficiency with MEMS steering. The combination of a reflective array of MEMS mirrors with a step steering approach can be very efficient over wide angles.

For **MEMS mirror arrays**, it is critical to have high area fill factors. One way to boost fill factors is to use hidden actuators, i.e., **micro-actuators** placed under mirror plates. Hidden actuators can be easily implemented in thin-film surface-micromachining processes.

MEMS mirror arrays need to have very high reflectivity in order to handle high powers because the thermal path for cooling is very slow.

Electrowetting Beam Steering for Lidar

Electrowetting has recently become another option for steering optical beams. The approach uses small boxes that contain **low-index water** and **high-index oil**. Each sidewall of the boxes has an electrode that modulates the liquid–surface-wetting contact angle. A prism geometry is created by applying sufficient voltage to any two opposing sidewalls to create a straight but tilted interface between the water and oil.



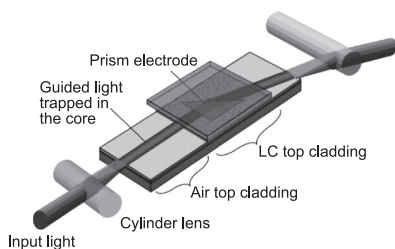
Electrowetting can easily change the apex angles of the prism by ± 45 deg to allow **transmissive beam steering** out to ± 45 deg. Compression and transmission ratios for the device can reduce net beam-steering efficiency to $\sim 50\%$ at the maximum steering angles.

Electrowetting is a fixed-period beam-steering approach. Steering occurs by changing the **order of the blaze**. If an additional phase layer is used, continuous steering can be achieved. The size of these elements can be up to a few millimeters.

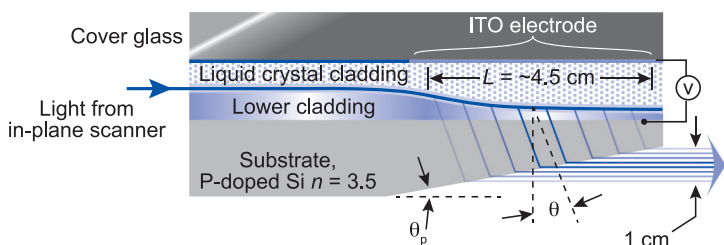
Electrowetting can be used in either **reflective or transmissive modes**. Switching speed should be on the order of 1 msec for the smallest prisms (tens of μm) and tens to hundreds of msec for millimeter-sized prisms. This size dependence stems from electrowetting being an electromechanical effect that requires movement of mass (i.e., the liquids). Optical absorption is practically a nonissue in the visible and NIR. For wavelengths $> 2 \mu\text{m}$, the water inside the prism is a strong absorber. To minimize optical absorption in these wavelengths, either the water must be replaced with a different liquid or the prisms must be very small (tens of μm high).

Steerable Electro-evanescent Optical Refractors

This refraction device uses LCs as an **active cladding layer** in a **waveguide architecture** where light is confined to a high-index core and the evanescent field extends into the variable-index LC cladding. This setup allows very large optical path delays (about 2 mm), so for small apertures, it eliminates the need for a 2π reset, and fringing fields are not an issue. Because the LC layer is thin, these devices can be relatively fast, with a $<500\text{-}\mu\text{sec}$ response time.



In-plane steering is accomplished by changing voltage on one or more prisms filled with LCs, as shown to the left. **Out-of-plane steering** is based on the **waveguide coupler** designed by R. Ulrich at Bell Labs in 1971, shown below.



In any waveguide, if the cladding is too thin, light leaks from the guided mode. In a planar slab waveguide, Snell's law gives the propagation angle of the escaping light; because the effective index of the waveguide is tunable, the angle of the escaping light is also tunable. This allows us to emit light at various out-of-plane angles.

This device can rapidly steer over relatively wide angles but is limited to small apertures. In-plane steering can be up to 60 deg or more in either direction, depending on the exact geometry. Out-of-plane steering can be ~ 15 deg in either direction. The main limitation on this technique is the size of the apertures, which must be ≤ 1 cm.

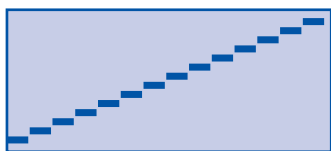
The development of this steering technique for short-range lidars will be interesting.

Electro-optical Effects

EO scanning, or lensing, uses either a **linear EO effect** (as in Pockels cells) or a **quadratic Kerr effect** and has the benefit of being fast. LiNbO_3 , GaAs, SBN, and BaTiO_3 are **linear EO crystals**, with LiNbO_3 being the most commonly used for an EO effect. SBN and BaTiO_3 can have large EO coefficients.

PMN-PT, PLZT, and KTN are nonlinear **Kerr-effect EO crystals**, in which the refractive index change is proportional to the square of the electric field. The variation in refractive index at 500 V/mm is 0.00009 for LiNbO_3 , 0.005 for BaTiO_3 , 0.00054 for PLZT, and 0.0152 for KTN; thus, **KTN** has the strongest Kerr effect and the strongest EO coefficient. This means either that a shorter interaction length can achieve the same deflection angle, or that less voltage and the same interaction length can be used.

KTN has only recently become available in crystals of appreciable size (multiple millimeters thick). The available beam steerers using KTN make use of charge injection with titanium electrodes, so a single electrode can be placed on top and bottom. This creates an effective lens that moves linearly with voltage to create angle steering.



Phase ramp across an EO crystal

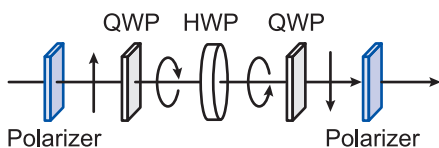


Electrodes on corners produce electronic lens

Placing a stair step of voltages across a fixed-width EO crystal should enable deflection of a laser beam. To compensate for the quadratic effect, the voltage change would need to be the square root across the crystal instead of being linear. Since these crystals require high voltage, it can be difficult to pattern the electrodes. If voltages are placed at the corners, a parabolic voltage change and a parabolic refractive index change can be obtained, making a writeable lens.

Polarization Birefringent Grating Beam Steering

A promising wide-angle steering approach is based on **circularly polarized gratings**. This approach directly creates a phase delay, eliminating the need for a 2π flyback region, thus, increasing steering efficiency. The steering is based on the optical model of the **quarter-wave plate, half-wave plate, quarter-wave plate (QHQP)** device of Pancharatnam. Linear incident light becomes circularly polarized after the first quarter-wave plate and can be defined as E_{in} according to Jones calculus notation:



quarter-wave plate, half-wave plate, quarter-wave plate (QHQP) device of Pancharatnam. Linear incident light becomes circularly polarized after the first quarter-wave plate and can be defined as E_{in} according to Jones calculus notation:

$$\mathbf{E}_{in} = \begin{bmatrix} Ex_{in} \\ Ey_{in} \end{bmatrix} = \begin{bmatrix} Ex_{in} \\ iEx_{in} \end{bmatrix}$$

For convenience, we assume right circularly polarized light. Ex_{in} and Ey_{in} are vector components along the x and y axes, respectively. The transmitted light E_{out} is defined as a linear mapping of the incident light E_{in} by a Jones matrix that represents the half-wave plate:

$$\mathbf{E}_{out} = \begin{bmatrix} \cos\beta & -\sin\beta \\ \sin\beta & \cos\beta \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 \\ 0 & e^{i\phi} \end{bmatrix} \cdot \begin{bmatrix} \cos\beta & \sin\beta \\ -\sin\beta & \cos\beta \end{bmatrix} \cdot \begin{bmatrix} Ex_{in} \\ iEx_{in} \end{bmatrix}$$

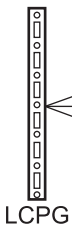
where β is the angle between the slow axis of the half-wave plate and the x axis, and ϕ is the phase retardation of the half-wave plate, which equals π . The final relationship can be simplified as

$$\mathbf{E}_{out} = \begin{bmatrix} Ex_{in} e^{i2\beta} \\ -iEx_{in} e^{i2\beta} \end{bmatrix}$$

Now the transmitted light is left circularly polarized with a common phase factor $e^{i2\beta}$. Its phase can be accurately controlled by the azimuth angle β . If β varies horizontally from 0 to π , the spatial phase profile of transmitted light varies horizontally from 0 to 2π . Using (1) a LC cell with an in-plane director, (2) an azimuth angle linearly rotating from 0 to π , and (3) a total OPD across the cell agreeing with the half-wave retardation for the design wavelength, the final spatial phase profile of transmitted light changes linearly from 0 to 2π . Although the optical thickness is only one half-wave, it is possible to create a constant large phase gradient over an aperture size that is limited only by manufacturing constraints.

Step Angle Steering with LC Polarization Gratings

Fixed **LC polarization gratings** (LCPGs) with nearly ideal diffraction efficiencies ($>99.5\%$) have been experimentally demonstrated over a wide range of grating periods, wavelengths, and areas. Each fixed LCPG stage can double the maximum steered angle in one dimension without major efficiency reductions, so very large steered angles are possible (to at least ± 40 -deg FOV). This is in contrast to volume holographic wide-angle steering, which requires fine angle steering both before and after the holographic stack of glass.

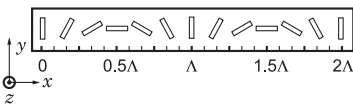


The figure shows **ternary steering**, which is possible if the LCPG can be turned off. The structure at the heart of these devices is a polarization grating that is implemented using nematic LCs. The nematic director is a continuous, in-plane, bend-splay pattern established using a UV polarization hologram exposing photo-alignment materials. When

voltage is applied, the director orients out of plane, effectively erasing the grating. Diffraction occurs according to

$$\eta_{m=0} = \cos^2\left(\frac{\pi\Delta nd}{\lambda}\right) \quad \text{and} \quad \eta_{m=\pm 1} = \left[\frac{1 \mp S'_3}{2}\right] \sin^2\left(\frac{\pi\Delta nd}{\lambda}\right)$$

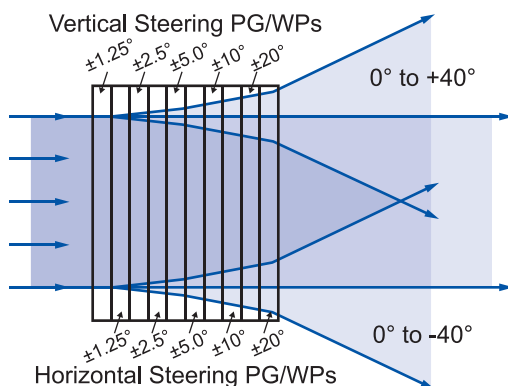
where η_m is the diffraction efficiency of the m^{th} order, λ is the wavelength of incident light, and $S'_3 = S_3/S_0$ is the normalized Stokes parameter corresponding to ellipticity of incident light. Only these three orders are possible, and when the retardation of the LC layer is a half-wave ($\Delta nd = \lambda/2$), 100% of the incident light can be directed out of the zeroth order. When the input polarization is circular, all of the light can be directed into a single first order, with the handedness ($S'_3 = \pm 1$) selecting the diffraction order.



The figure shows **rotating phase**, which has no resets.

Multiple-Stage LCPGs

Multiple stages can be cascaded to implement **digital beam steering**, as shown in the figure, where PG/PW is polarization grating / wave plate. The number of stages within the beam-steering stack limits the overall efficiency, which is strongly affected by four factors: steering efficiency, Fresnel loss, absorption, and scatter.



Steering loss is the fraction of received light not steered into the intended order. Current loss is $<0.5\%$ per grating. **Fresnel loss** is due to refractive index differences; the initial loss entering a stack is higher

than that within a stack because of index matching in the stack. **Absorption** is primarily due to the transparent conductor layers if indium tin oxide is used; newer transparent conductors can limit this loss. Current loss from **scatter** is $<0.3\%$ for the best passive gratings.

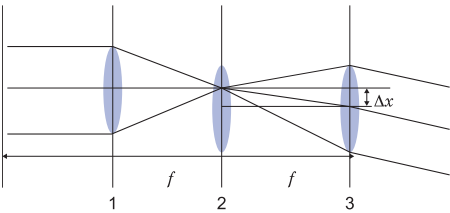
Using novel schemes, up to 87% efficiency over an 80×80 -deg steering FOV is possible, with 1.25-deg steering increments. With careful engineering, this approach combined with narrow-FOV LC steering should be able to achieve $>90\%$ efficiency over a very wide FOV.

Speed is limited by the ability to switch polarization. A polarization rotation stage is included between each layer. Room temperature nematic LCs switch at ~ 10 msec. Polymer network LCs can increase the switch speed but can also suffer from additional scattering loss.

Ferroelectric LCs (FLCs) also increase the switch speed, probably without additional scattering loss. With FLC polarization, rotation switching speeds can be reduced to ~ 50 μ sec.

Lenslet-based Beam Steering

A **microlens array** uses three cascaded arrays to steer a beam to discrete locations. A drawback of **lenslet-based beam steering** is that it uses multiple layers. Each layer adds complexity and can reduce steering efficiency due to engineering factors such as absorption, reflection, and scattering.



A single triplet of the cascaded arrays is shown.

Early designs used a doublet for steering, but it was found that a field lens

between the two lenslets redirects light such that almost all of the light passing through the first lens reaches the last lens. The **steering angle** is given by

$$\tan \theta = \frac{\Delta x}{f_1}$$

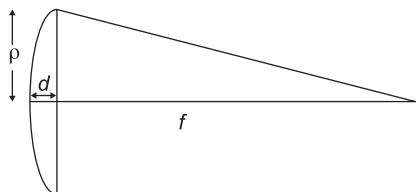
where Δx is the amount of decenter of the first microlens with respect to the optical axis of the other two microlenses, and f_1 is the common focal length of all three microlenses in the triplet. Steering is based on changing the diffractive orders, so **phase corrector plates** dither the angular location of orders or, alternatively, a fine-angle beam steerer steers between orders.

An example of **steering to discrete angles** is shown in the table. At $f/1$, steering up to 25 deg can be achieved; thus, this is a moderately **wide-angle steering** approach, as well as an approach that is moving toward high efficiency.

Δx [μm]	Steering angle [deg]	Focal length [μm]	Lenslet diameter [μm]	$f/\#$	OPD [μm]
67.5	25	145	150	1	19.4
67.5	20	185	150	1.2	15.2
67.5	10	383	150	2.6	7.3
67.5	3	1288	150	8.6	2.2
67.5	1.5	2578	150	17.2	1.1

Electronically Written Lenslets

Creating a moving microlens array electronically allows for very precise positioning of the array because the array positions are based on electrode spacing and applied voltages. For simplicity, plano-convex microlens elements



are assumed, as illustrated. The lens' required maximum thickness d for creating the OPD can be found using the standard **sag formula**:

$$OPD = (n - 1)d = \frac{\rho^2}{2f_1}$$

where n is the refractive index of the microlens, and f_1 is the focal length of the lens. If $\Delta x = \rho$, the radius of the microlens, the lenslet array will steer to the maximum allowed angle. Using $\Delta x = \rho$,

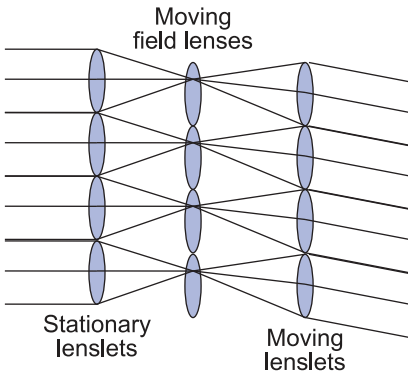
$$OPD = \frac{\rho}{2} \tan(\theta_{\max})$$

The effective resets are at the spacing of the lenslet arrays. However, no physical reset is involved, so no fringing-field issues arise. A change in slope of the OPD occurs at the inflection point between lenslets, but no discontinuity occurs.

The OPD must be $4\times$ thinner than a sawtooth profile phase in order to steer to the maximum angle in a lenslet array. A factor of 3.5 reduction in OPD using this approach is a good estimate.

The most effective method of fine-angle beam steering between orders involves adding (up to) a one-wavelength phase delay behind each lenslet in one of the lenslet arrays. This additional **phase delay** dithers the location of the orders of the fixed-period beam steerer. The additional phase delay can be included in one of the lenslet arrays or can be a separate, fourth layer, depending on which configuration is easier to implement.

Mixed-Lenslet Arrays



Another lenslet-based beam-steering approach uses a **mixed set of lenslet arrays** (as shown), with both positive and negative lenslets. This approach eliminates the foci, thus increasing power-handling capability. A second major advantage of the mixed-lenslet beam-

steering approach is its ability to use lenslets that are too small to have a quality focus. Small lenslets with a high F-number do not have a quality focus, but this size effect does not significantly degrade the optical system in a mixed-lenslet array.

The focal length of the positive lenslets is twice that of the lenses in the **all-positive lenslet approach** if the steering angle is kept constant, while the negative lenslets in the mixed-lenslet case have a focal length half as large as that in the all-positive lenslet approach but can be implemented using a high-index-fixed material.

This approach requires the two positive outside lenslet arrays to have identical focal lengths and to be placed one focal length apart. For **mixed-lenslet steering**, the center lenslet is moved down one unit, while the positive lenslet on the other side is moved up two units. For a mixed-lens approach,

$$\theta = \arctan\left(\frac{2\Delta x}{f_1}\right)$$

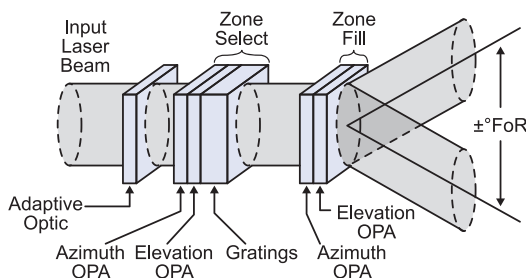
where f_1 is the focal length of the positive lenslets. The focal length of the negative lenslet is one-quarter that of the positive lenslet.

Digital simulations show the mixed-lenslet approach having >98% steering efficiency for small angles such as 1.5 deg. The middle lenslet can still clip energy entering the first lenslet, but the main efficiency issue is loss through multiple layers.

Holographic Gratings for Beam Steering

Volume (thick) holograms offer the potential to implement **large-angle steering** with high efficiency. Using multiple holograms, multiple discrete steering angles can be addressed. A small-angle beam steerer selects the hologram to be addressed, resulting in the desired large angle of deflection. Another small-angle continuous-beam steerer is then required to go between the large angles addressed by each hologram.

The figure shows an adaptive optics layer followed by azimuth (az) and elevation (el) small-angle beam steerers, then a stack of volume holograms followed by another set of az and el beam steerers.



Advances in **volume holographic gratings** in photo-thermal refractive glass make for an interesting large-angle step-steering approach. Each glass holographic grating can handle high power and can have >99% efficiency. When two holograms are written in a single piece of glass, efficiency is still over 98%. Losses due to scattering and absorption are less than 0.5% for each grating. Currently, the largest holographic glass available is about 5 cm in diameter.

Many layers of holographic glass can be placed back to back with low loss. For example, with 8 holograms placed in each direction (az and el), if each hologram steers to an angle separated by 5 deg from the adjacent angle, we have a total FOV of 40 deg, broken into 8 zones of 5 deg each.

The biggest disadvantage of volume holograms for wide-angle step steering is that another hologram is required for each new set of wide angles being addressed.

Geometrical Optics

Geometrical optics is an important subject for lidar. One topic in geometric optics that is central to lidar is **long-focal-length optics** to provide high resolution at long range for object identification. Lidar is often cued by an acquisition sensor, and the lidar's task is to identify an object. To accomplish this, as much resolution as possible is required. Also, it is likely that the acquisition sensor is a thermal imager at longer wavelengths than the lidar, so its diffraction-limited resolution will not be as good as the lidar's at, for instance, $\sim 1.5\text{-}\mu\text{m}$ wavelength. A key equation for **focal spot diameter** is

$$D = 2(f/\#)\lambda$$

where the factor of 2 captures most of the energy in the spot, and F-number is focal length over aperture diameter:

$$f/\# = \frac{f}{D}$$

It is clear that the wavelength and the F-number, not the size of the optical aperture, are important. While for a small F-number the focal spot can be about one wavelength, a large F-number is often needed in order to have sufficiently large detectors to sample the incoming field as accurately as needed.

The **detector angle subtense** (DAS) $\times$ focal length is often set approximately equal to the spot diameter, although oversampling the smallest spot diameter can also be of interest. Another term often used in lidar is the **numerical aperture** (NA), given by

$$NA = \frac{1}{2(f/\#)}$$

The **focusing ability** is important if stops are used to block radiation outside of a certain region. An intermediate focus can be used on receive with solid material around it to block any radiation much larger than the spot size.

Inertial Measurement Units

A lidar image is often made by stitching together sections of the image gathered at various times. An **inertial measurement unit** (IMU) is a device that measures and reports on a lidar's velocity and orientation using a combination of **accelerometers** and **gyroscopes**. The lidar will possibly use an IMU on a vehicle; however, when measuring precise changes in its position or velocity, it is more likely that the lidar will have its own IMU. Unless the vehicle is perfectly rigid, an IMU in one location will not record exactly the same motion as an IMU in another location.

A flash image gathered using a LMAPD can generate a complete image with a single frame, but all of the other types of lidar require **stitching** together multiple samples taken at different times for which an IMU is critical. A **flash image** generated using a **GMAPD** is not truly a single image. As many as, say, 50 images might be taken of the same target to create one flash image, so motion between samples must be compensated for.

With a laser repetition rate of 40 KHz, frames need to be stitched over about 1 msec. A scanning imager collects data at various times, so motion must be compensated for. SAL also requires motion compensation.



An IMU works by detecting the current rate of acceleration using one or more accelerometers and detects

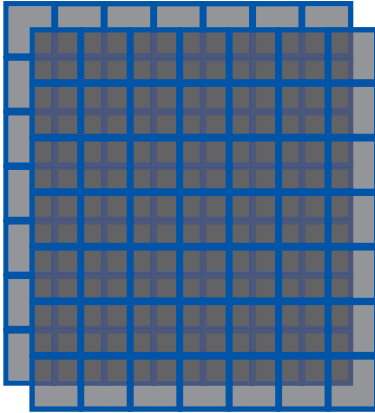
changes in rotational attributes such as **pitch**, **roll**, and **yaw** using one or more gyroscopes.

IMUs are used for lidar navigation but need to be periodically reset. If they are being used to stitch together lidar images in which portions of the image were gathered at different times, **relative motion** might be sufficient. If **absolute motion** is required, IMU data might need to be supplemented with GPS or other navigation information.

Microscanning of Lidar Images for Improved Sampling

Sometimes it is useful to sample at a higher **sampling density** than is provided by the DAS. To this end, we can

Image 1 ↘



↗ Image 2 misregistered by $\frac{1}{2}$ DAS

take an image with the FPA in one location and move the FPA a fraction of the DAS (often half of the DAS) to make a new image, as shown in the figure.

This method increases sampling in angle/angle space and can be used with passive sensors as well as with lidars. However, a method to impose the small motion is needed.

An ideal method of **imposing the small motion** uses a lidar that is not well stabilized combined with an IMU that provides the lidar location. This is called **random microscanning**, unlike microscanning that is imposed by uniform motion.

Microscanning can interact with the various beam-steering approaches. The entire FPA can be stepped up from one location to another to increase sampling, or **micromotion** can be imposed. Both methods increase sampling to the same degree, but micromotion requires much less angular motion.

Another consideration when microscanning is the size of the **focal spot**. Resolution can be limited by the DAS or by the size of the focal spot. Microscanning increases sampling but does not increase resolution, so both the DAS and the diameter of the focal spot are larger than the amount of microscan micromotion. This limits the focal lengths that can be considered for use in lidar design.

Range Measurement Processing

A lidar measures range by taking the convolution of the “pulse” it emits with the range profile of the target. This creates a **measured range profile** of the target, corrupted by the fact that the lidar pulse is not a delta function. The range “to the target” is actually a measured range profile.

A simple rectangular pulse can be assumed to have a certain pulse width, a Gaussian pulse shape, or another more complex pulse shape. A linear frequency chirp or a pseudo-random-coded waveform will have an equivalent pulse shape.

A **convolution** is given by

$$f * g = \int_{-\infty}^{\infty} f(\tau)g(t - \tau)d\tau$$

This convolution can also be performed digitally. $\text{Conv}(u,v)$, where u and v are the two range profiles in this case, is a MATLAB[®] command to perform a digital convolution. The above equation can be made discrete for a **digital convolution**, becoming

$$w(k) = \sum_j u(j)v(k - j + 1)$$

A processing issue with respect to **range measurement accuracy** is the **lidar clock frequency**. If the receiver is digitally sampled, this sampling rate limits the range resolution. For example, a digital clock frequency of 0.25 nsec works well for sampling a 3-sec pulse but would not work for sampling a 0.5-nsec pulse. If the **clock frequency drifts** during the processing, an error occurs because it is assumed that the signal is sampled every 0.25 nsec. Any deviation in the clock frequency causes an error.

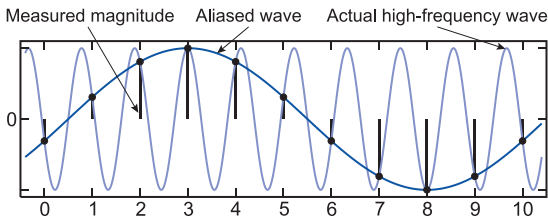
The **matched filter** of an object is the autocorrelation of the target with an impulse response. Therefore, performing ideal processing on a delta function waveform results in a matched filter.

Nyquist Sampling a Range Profile

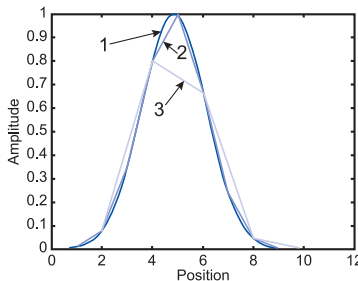
The **Nyquist sampling theorem** states that a signal can be reconstructed from a set of samples if it is sampled at at least twice its maximum frequency.

In order to **reconstruct the range profile** of a target, the **Nyquist criterion** must be met; i.e., to accurately reconstruct the shape of the return range profile, one must sample more frequently than twice during the highest frequency. If sampling is not performed frequently enough, aliasing can occur, in which a high frequency is aliased down to an incorrect lower frequency.

Sampling can be performed in Fourier space or in real space. In Fourier space, the return range profile pulse is converted into a **series of sine waves**. The first figure shows how the same data can represent more than one sine wave possibility. If a sine wave is not sampled often enough, the result appears to be a lower frequency than it actually is.



The second figure plots (1) a **Gaussian-shaped pulse**, (2) that same pulse sampled once every unit on the x scale, and (3) that same pulse sampled once every other unit on the x scale. When sampling is reduced, the accuracy of replicating a Gaussian-shaped pulse is also reduced.

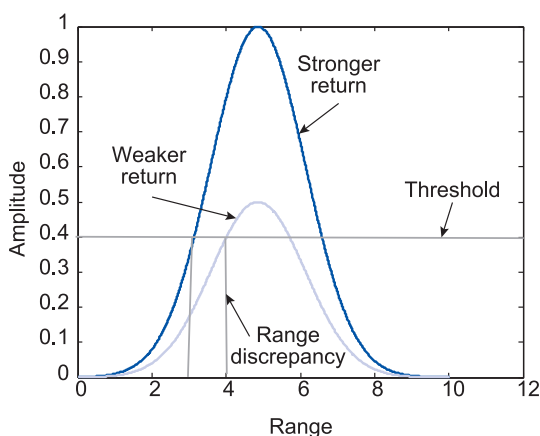


Threshold, Leading Edge, and Peak Detectors

Some detectors, such as GMAPDs, use the entire pulse to sample range. This is necessary because in this case there is some probability of an avalanche occurring any time during the pulse. Other detectors can use the leading edge (or peak) of a pulse to determine range.

Leading edge or **threshold detectors** use only a portion of the pulse for range measurement. These detectors also provide higher range accuracy for a given pulse width compared to those that sample range using the entire pulse. In theory, the pulse width does not matter if the leading edge alone is used.

Using only a portion of the pulse also requires more laser power. **Leading edge detection** can be biased by higher-versus-lower SNR, as can be seen in the figure. The object is in the same range, but the higher SNR shows it closer.



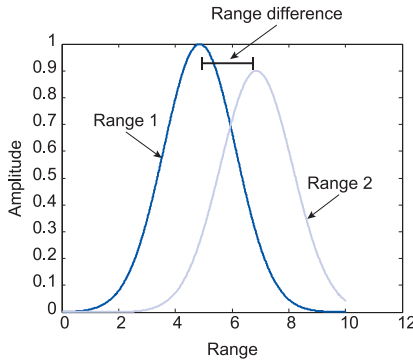
If range can always be measured as a ratio of pulse amplitudes, always measuring at, for instance, the initial 40% or 50% level, this bias can be eliminated; however, both the peak and the threshold levels must be measured. The figure shows that the 50% value from the peak is the same for both signal amplitudes. For this reason, this is sometimes called **constant fraction detection**.

Peak detection by itself is somewhat unreliable because any noise causes fluctuations in the peak value.

Range Resolution, Precision, and Accuracy

The terms range resolution, range precision, and range accuracy are often used incorrectly or carelessly, resulting in unnecessary confusion.

Range resolution quantifies the ability to detect two objects separated in range along the same line of sight, as shown in the figure.



Range resolution is the distance by which two range profiles need to be separated such that the sum of the intensities has a dip between the two range intensity profiles. The bandwidth of the signal determines the range resolution, which is limited to $c/(2B)$.

To make use of range resolution, multiple range measurements need to be made. For LMAPDs, this means using a ROIC that can capture multiple range measurements.

Range precision quantifies the relative uncertainty of a measurement of range to an object. High SNR allows for measurement of range precision to a small fraction of the range resolution. The better the SNR the higher the ratio of range precision to range resolution. This is a concept that certainly causes confusion.

Range accuracy depends on range precision as well as systematic errors (e.g., clock rate, drift, timing offsets, etc.). It quantifies the degree to which a range measurement yields the absolute value, or true value, of range.

Fourier Transforms

In terms of computation, the difference between the pupil plane and the image plane is a **Fourier transform**. Alternatively, this difference can be a lens. Once the field is known in either domain, the other domain can be known.

The aperture receives the pupil plane field, and a lens focuses that field into the image plane. Once in the image plane, either the full field can be captured using heterodyne techniques, or intensity alone can be captured, as in direct detection.

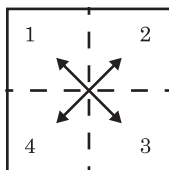
The full field in the pupil plane can be captured using a Fourier transform instead of a lens, in what is termed **lensless imaging**. The Fourier transform of a pupil plane field is given by

$$F\{E(x,y)\} = \iint E(x,y)e^{[-i2\pi(f_x x + f_y y)]} dx dy$$

This equation transforms the electric field into an image plane, spatial-frequency-based complex term that contains the far field image information. The translation from the far field object plane to the pupil is a Fourier transform of the far field object plane to obtain the pupil plane field. We use a lens or a Fourier transform to move from the pupil plane to the image plane, recovering the object plane field in the image plane.

Most engineers perform Fourier transforms digitally, using the **fast Fourier transform** (FFT) and MATLAB[®] commands, examples of which are listed here:

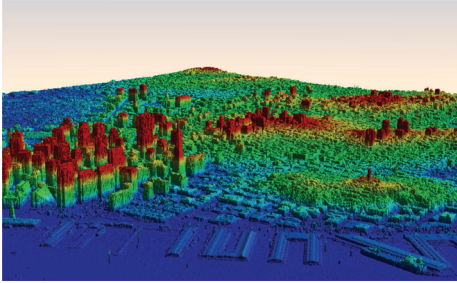
- $Y = \text{fft2}(X)$ returns a 2D digital FFT of a 2D variable X .
- $Y = \text{fft2}(X,m,n)$ pads the area around X with zeros to enlarge the Fourier transform.
- $Y = \text{fft}(X)$ returns an array that is the same size as the original array. (To provide more fidelity in frequency space, sometimes the returned array needs to be larger than the original.)



- $Y = \text{fftshift}$ moves the zero-frequency components to the center of the array; it swaps the locations of the terms in a 2D array, as shown.
- $Y = \text{fftshift}(\text{fft2}(X,m,n))$ both pads the array and shifts the center.

Developing 3D Maps from Lidar

Angle/angle information can be obtained by scanning lidar or flash lidar, both of which can use GMAPDs or LMAPDs. All lidars obtain range by sending out a waveform and measuring the time required for the waveform to return. The waveform



can be a series of pulses or some other waveform, but time of flight is measured and the speed of light is used to determine range to the object. This range measurement occurs in each angle/

angle pixel, creating a **3D voxel**. The figure shows a 3D image of San Francisco.

Scanning lidars require rapid movement in angle. For example, assume that 8 individual detectors need to make an image that is 480×640 pixels in angle space. We could have 60 rows of 8 pixels times 640 pixels in each row. At 30-Hz frame rate, we would need to move $60 \times 640 \times 30 = 1,152,000$ times per second and have a laser with that repetition rate. It would take 0.033 sec to make an image, so we would need to determine how far the lidar and target moved during that 0.033 sec.

We could also make the same image with a 32×128 -pixel **GMAPD camera**. An array of 15 snapshots high would provide the 480 pixels by 5 snapshots wide to provide the 640 pixels. Assuming 50 pulses per snapshot, we would need a laser repetition rate of $50 \times 5 \times 15 \times 30 = 112,500$ times per second. We would need to stabilize only to a fraction of the 32×128 array; stitching together the frames in the 128×32 section would occur in a little over 1 msec. It will take 0.033 sec to make that image.

A third option is to use the same 32×128 array and make the image with a **LMAPD**. The laser would need to pulse at a rate of $5 \times 15 \times 30 = 2250$ Hz. If the LMAPD array was 480×640 , the laser would only need to pulse at 30 Hz.

3D Metrics for Lidar Images

The goal of using a sensor is to detect, recognize, or identify an object. The sensor parameters needed to obtain the desired information depend on the object of interest. For example, recognizing a person's face requires only a small FOV, but the resolution required is on the scale of centimeters. A military tank is much larger, and its features are not as fine, so a larger FOV is required, but resolution can be relaxed to, say, 15 cm.

The following standards have been established for metrics used in imaging, some of which can be adapted to 3D lidar imaging:

- **Johnson's criteria** relates the target-dependent aspect of sensor-based detection. Presented in 1958 in a paper entitled "Analysis of image forming systems," by John Johnson (U.S. Army Research and Development Laboratories), this is the first set of sensing metrics to be widely used.
- The **National Imagery Interpretability Rating Scale (NIIRS)** was made available in the 1990s. This scale is based on image interpretability for a large variety of tasks as a method to evaluate the relative value of imagery. No NIIRS metric currently exists for 3D lidar, but it is anticipated that one will be developed. A NIIRS metric for 3D lidar will allow comparison between various lidars.
- A **general image quality equation (GIQE)** was recently developed for visible and NIR imagery but has not yet been developed for 3D lidar imagery. The GIQE predicts NIIRS ratings.

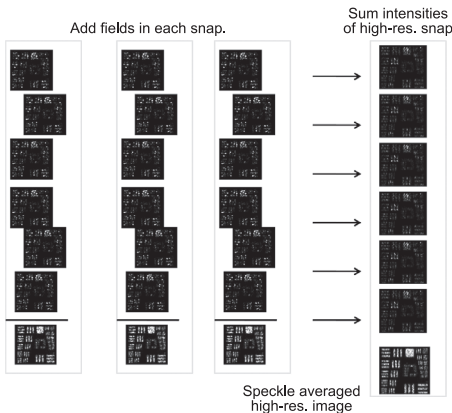
Multiple-Subaperture Spatial Heterodyne Processing

Spatial heterodyne detection measures the field across a FPA in either the image plane or the pupil plane. If the field is measured in the pupil plane and is Fourier transformed, a bright central area results. This area is the **autocorrelation** of the LO summed with the autocorrelation of the signal, which has much a lower intensity than the LO:

$$I = |u_t + u_{LO}|^2 = |u_t|^2 + |u_{LO}|^2 + u_t u_{LO}^* + u_t^* u_{LO}$$

The two first terms in this equation are in the center of the field, but the LO intensity dominates. The remaining two terms are spatially separated and are complex conjugates of each other. The complex terms provide field information as well as intensity.

The next step is to gather multiple “snaps” or images and transform each to the image plane. The motion of each subaperture is not known, although an inertial system can provide an indication of motion. The locations are dithered one subaperture snap at a time, and one image quality metric, such as image sharpness, is considered. With **multiple subapertures**, moving only one subaperture does not significantly change the sum metric. For each snap, the optimum location for best image quality is determined by summing the dithered subapertures.



Once the subaperture location is determined in each snap, we sum the snaps for each subaperture and add them to determine the speckle-averaged image, as shown in the figure.

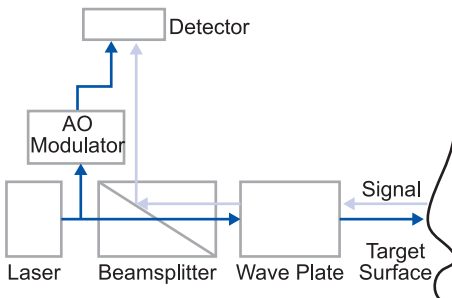
Definitions of Lidar Data Processing Stages

The following definitions of **lidar data processing stages** are used by the National Geospatial Agency (NGA) in the United States and many other organizations but are not necessarily used by the entire lidar community.

- **Level 0 (L0) – Raw data and metadata:** L0 data consists of the raw data in the form in which it was stored when collected from the mapping platform.
- **Level 1 (L1) – Unfiltered 3D point cloud:** L1 data consists of a 3D point data (point cloud) representation of the objects measured by the lidar mapping system. It is the result of applying algorithms to project the L0 data into 3D space.
- **Level 2 (L2) – Noise-filtered 3D point cloud:** L2 data differs from L1 in that noisy, spurious data have been filtered from the dataset, intensity values have been determined for each 3D point, and relative registration has been performed. The impetus behind separating L1 from L2 is due to the nature of **Geiger-mode lidar** (GML) data. Derivation of L1 GML data produces very noisy point clouds and requires specialized (coincidence) processing to remove the noise.
- **Level 3 (L3) – Georegistered 3D point cloud:** L3 datasets differ from L2 datasets in that the L3 data have been registered to a known geodetic datum.
- **Level 4 (L4) – Derived products:** L4 datasets represent lidar-derived products to be disseminated to standard users. These products could include digital elevation models (DEMs), viewsheds, or other products created in a standard format and using a standard set of tools.
- **Level 5 (L5) – Intel products:** L5 datasets are a type of specialized products for users in the intelligence community. These products can require specialized tools and knowledge to generate.

Processing Laser Vibrometry Data

Laser vibrometry generates a signal in frequency space. One way to measure that signal is shown in the figure.



High Doppler shift implies high velocity between the object and sensor.

Laser vibrometer signal processing involves: (1) highlighting specific frequencies contained in the return signal,

(2) converting the frequencies to velocities, and (3) converting the velocities to vibration spectra. Measuring a certain frequency takes time. The Doppler shift is

$$\Delta v = \frac{2V}{\lambda}$$

Measuring the **Doppler shift** requires at least one over twice the Doppler shift:

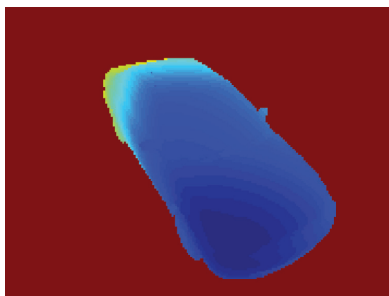
$$\Delta t = \frac{2}{\Delta V} = \frac{\lambda}{V}$$

To measure a Doppler shift of 2 Hz, one would need to stare at a given location for a full second. To measure low-frequency Doppler shifts (micro-Doppler velocities), one would need to stare at a given location for an even longer time. This spectrogram processing approach begins to break down when the Doppler frequencies become comparable to the vibrational frequencies.

Gathering vibration data sequentially is not the same as gathering data simultaneously, as in 3D imaging, where only lidar motion needs to be accounted for. When measuring vibration, the object is moving with time. Ideally, one could view the entire modal structure of an object (a bridge, for instance) at the same time. It would be difficult, if not impossible, to reconstruct the modal structure one piece at a time.

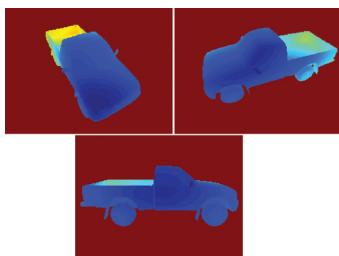
Target Classification Using Lidar

3D shape is an ideal characteristic to use for **object identification** because, in general, it does not change significantly. Of course, someone can pile brush on top of a vehicle to disguise its shape; however, the basic shape of an object is much less variable than the grayscale of the 2D image illuminated with variable light sources, which might even illuminate only a portion of an object.



3D point cloud images are rendered in terms of **voxels**, which define a point in 3D space, rather than in pixels, which define a point in 2D space. The figure shows a 3D image rendered with color showing range.

It is usual for the lidar range dimension to be as high, or higher, in resolution compared to the angle/angle dimensions because the range resolution depends on the lidar bandwidth, not on the diffraction limit. At long distances, it is therefore easy to have excellent range resolution with a lidar. This is in contrast to a stereo 3D image, where the range resolution will not be as good as the angle/angle resolution.



One useful advantage of a 3D point cloud image is that the image can be viewed from multiple angles, as shown. While data from angles other than the angle viewed by the lidar might be missing, often there is sufficient data to provide excellent 3D views from many different

locations. Not only does this make it much easier to hand off from one sensor to another, it also improves the ability to recognize an object. An object is often more recognizable from one particular aspect than from another aspect.

Equation Summary

Range resolution:

$$\Delta R = \frac{c}{2B}$$

Current in a heterodyne lidar:

$$I = 2E_{\text{sig}} \cdot E_{\text{LO}} \cdot \exp[-jt(\omega_{\text{sig}} - \omega_{\text{LO}})]$$

The full width, half maximum diffraction limit:

$$\vartheta \approx \frac{1.03\lambda}{D}$$

Gaussian shape:

$$f(x) = ae^{-\frac{(x-b)^N}{2c^2}}$$

Coherence length:

$$C_l = \frac{c}{B}$$

Laser linewidth in wavelength:

$$B = \frac{c}{\lambda^2} \Delta\lambda$$

Power received:

$$P_s = P_T \cdot \frac{\sigma}{A_{\text{illum}}} \cdot \frac{A_{\text{rec}}}{\pi R^2} \cdot \eta_{\text{atm}}^2 \cdot \eta_{\text{sys}}$$

Optical cross section for a Lambertian scatterer:

$$\sigma = \rho_t A_p = \rho_t d^2$$

Optical cross section for a trihedral corner cube:

$$\sigma \approx \frac{4a^4}{\lambda^2}$$

Equation Summary

Wavelength change for a new speckle realization:

$$\Delta\lambda = \frac{\lambda^2}{\Delta z - \lambda}$$

Exponential decay (Beer's law):

$$E = E_0 e^{-(\beta R)}$$

Coherence time:

$$\tau_0 = \frac{r_0}{\nu}$$

Transmitted energy:

$$E_T = E_R \cdot \frac{A_{\text{illum}}}{\sigma} \cdot \frac{\pi R^2}{A_{\text{rec}}} \cdot \frac{1}{\eta_{\text{atm}}^2} \cdot \frac{1}{\eta_{\text{sys}}}$$

Signal-to-noise ratio:

$$SNR = \frac{\langle i_s^2 \rangle}{\langle i_n^2 \rangle}$$

Signal current in the detector:

$$\langle i_s^2 \rangle = \begin{cases} G^2 \Re^2 P_s^2 & \text{(direct detection)} \\ 2\eta_h G^2 \Re^2 P_{LO} P_S & \text{(coherent detection)} \end{cases}$$

Responsivity of the detector:

$$\Re = \frac{\eta_q e}{h\nu}$$

Noise current in the detector:

$$\langle i_n^2 \rangle = \langle i_{\text{shot,sig}}^2 \rangle + \langle i_{\text{shotLO}}^2 \rangle + \langle i_{\text{bk}}^2 \rangle + \langle i_{\text{dk}}^2 \rangle + \langle i_{\text{th}}^2 \rangle$$

Equation Summary

Ideal power received for direct detection:

$$P_s = \frac{Nhc}{\lambda T_m}$$

Signal-to-noise ratio for direct detection:

$$SNR = \frac{\langle i_s^2 \rangle}{\langle i_n^2 \rangle} = \frac{G^2 \Re^2 P_s^2}{\langle i_{\text{shot,sig}}^2 \rangle + \langle i_{\text{shotLO}}^2 \rangle + \langle i_{\text{bk}}^2 \rangle + \langle i_{\text{dk}}^2 \rangle + \langle i_{\text{th}}^2 \rangle}$$

Discrete Poisson probability density function:

$$q(k, M) = \frac{M^k e^{-M}}{k!}$$

Gaussian probability density function:

$$p(k) \approx \frac{1}{\sqrt{2\pi M}} e^{-(k-M)^2/2M^2}$$

Thermal noise:

$$E_{\text{th}} = kTB$$

Mean-squared thermal noise current:

$$\langle i_{\text{th}}^2 \rangle = \frac{4kTB}{R_L}$$

Thermal noise [in decibels above a milliwatt (dBm)]:

$$P_{\text{thdbm}} = 10 \log_{10}(1000kTB)$$

Energy per photon:

$$E_p = h\nu$$

Equation Summary

Energy of a photon:

$$E_p = \frac{hc}{\lambda}$$

Signal-to-noise ratio versus number of events:

$$SNR = \frac{M}{\sqrt{M}} = \sqrt{M}$$

Shot noise current:

$$\langle i_{\text{shot,sig}}^2 \rangle = 2eG^2F\Re P_s B$$

Radiance of the sun:

$$L_\lambda = \frac{2hc^2}{\lambda^5} \left[\exp\left(\frac{hc}{\lambda kT}\right) - 1 \right]^{-1}$$

Mean-squared background current:

$$\langle i_{\text{bk}}^2 \rangle = 2eG^2F\Re P_{\text{bk}} B$$

Received optical power:

$$P_R = \frac{Nhc}{\lambda T_m}$$

Mean-squared dark current:

$$\langle i_{\text{dk}}^2 \rangle = 2eB(I_{\text{dks}} + I_{\text{dkb}}G^2F)$$

Signal-to-noise ratio for direct detection:

$$\begin{aligned} SNR &= \frac{\langle i_s^2 \rangle}{\langle i_n^2 \rangle} \\ &= \frac{G^2 \Re^2 P_s^2}{2eBG^2F[\Re(P_s + P_{\text{LO}} + P_{\text{bk}}) + P_{\text{dks}}] + 2eBi_{\text{dks}} + 4(kTB/R_L)} \end{aligned}$$

Equation Summary

Background-limited SNR for direct detection:

$$SNR = \frac{\eta_D P_s^2}{2h\nu B P_{bk}}$$

Shot-noise-limited SNR for direct detection:

$$SNR_{\text{shotlim}} = \frac{\eta_D P_R}{2h\nu B}$$

Probability of M coincident detections out of N pulses:

$$P(M \text{ out of } N \text{ pulses}) = \frac{N!}{M!(N-M)!} (P_o)^M \times (1-P)^{N-M}$$

Signal electric field:

$$\vec{E}_{\text{sig}} = E_{\text{sig}} \exp(k_{\text{sig}} \vec{r}_{\text{sig}} - j\omega_{\text{sig}} t)$$

Local oscillator electric field:

$$\vec{E}_{\text{LO}} = E_{\text{LO}} \exp(k_{\text{LO}} \vec{r}_{\text{LO}} - j\omega_{\text{LO}} t)$$

Heterodyne detection current:

$$I = \vec{E}_{\text{sig}} \cdot \widehat{\vec{E}}_{\text{sig}} + \vec{E}_{\text{LO}} \cdot \widehat{\vec{E}}_{\text{LO}} + \vec{E}_{\text{sig}} \cdot \widehat{\vec{E}}_{\text{LO}} + \vec{E}_{\text{LO}} \cdot \widehat{\vec{E}}_{\text{sig}}$$

Signal-to-noise ratio for heterodyne detection:

$$SNR = \frac{\langle i_s^2 \rangle}{\langle i_n^2 \rangle} = \frac{\eta_h G^2 \Re^2 P_{\text{LO}} P_s}{eBG^2 F [\Re(P_s + P_{\text{LO}} + P_{\text{bk}}) + P_{\text{dks}}] + 2eB i_{\text{dks}} + 4(kTB/R_L)}$$

Shot-noise-limited SNR for heterodyne detection:

$$SNR = \frac{\eta_D \eta_h P_s}{h\nu B}$$

Temporal heterodyne signal electric field:

$$\vec{E}_{\text{sig}} = E_{\text{sig}} \exp(-j\omega_{\text{sig}} t)$$

Equation Summary

Temporal heterodyne local oscillator electric field:

$$\vec{E}_{LO} = E_{LO} \exp(-j\omega_{LO}t)$$

Beat current for temporal heterodyne detection:

$$I = 2E_{sig} \cdot E_{LO} \cdot \exp[-jt(\omega_{sig} - \omega_{LO})]$$

Beat current for spatial heterodyne detection:

$$I = 2E_{sig} \cdot E_{LO} \cdot \exp(k_{sig} \vec{r}_{sig} - k_{LO} \vec{r}_{LO})$$

Spatial heterodyne signal electric field:

$$\vec{E}_{sig} = E_{sig} \exp(k_{sig} \vec{r}_{sig})$$

Spatial heterodyne local oscillator electric field:

$$\vec{E}_{LO} = E_{LO} \exp(k_{LO} \vec{r}_{LO})$$

Shot-noise-limited SNR for spatial heterodyne detection:

$$SNR = \frac{\eta_D \eta_h P_s}{hfB}$$

Doppler shift:

$$\Delta f = \frac{2V}{\lambda}$$

Angular resolution for synthetic aperture lidar:

$$\Delta \vartheta = \frac{\lambda}{2L}$$

Equivalent angular motion using interferometric synthetic aperture lidar rotation:

$$\Delta \phi \approx \frac{L}{R}$$

Equation Summary

Range/Doppler velocity toward the target:

$$V_T = V \cos(\vartheta) \cdot \cos(\phi)$$

Optical parametric oscillator wavelength change:

$$\frac{1}{\lambda_p} = \frac{1}{\lambda_s} + \frac{1}{\lambda_i}$$

Mode-locked pulse separation:

$$\tau_m = \frac{2L}{c}$$

Mode-locked pulse width:

$$\Delta t \approx \frac{0.44}{P \Delta f}$$

Doppler shift velocity resolution:

$$\Delta V = \frac{\lambda^* \Delta f}{2}$$

Deflection angle for acousto-optical detection:

$$\sin \vartheta = \frac{m \lambda}{2 \Lambda}$$

Unambiguous range:

$$R_{\text{unambig}} = \frac{c \tau}{2}$$

Lens phase profile:

$$t_{\text{lens}}(w_{\text{az}}, w_{\text{el}}) = e^{\frac{-j\pi(w_{\text{az}}^2 + w_{\text{el}}^2)}{\lambda f}}$$

Diameter of the Airy disk to the zeros:

$$D_{\text{Airy}} = 2.44 \frac{\lambda f}{D}$$

Equation Summary

Beam-steering angle through a telescope:

$$\vartheta_f = \frac{\vartheta}{M}$$

Full array width:

$$D = nd$$

Relaxation time for a liquid crystal cell:

$$\tau_d = \frac{t^2 \gamma}{k \pi^2}$$

Turn-on time for a liquid crystal cell:

$$\tau_r = \frac{\tau_d}{\left[\left(\frac{V}{V_t} \right)^2 - 1 \right]}$$

Fringing-field-limited beam steering efficiency:

$$\eta = \left(1 - \frac{\Lambda_F}{\Lambda} \right)^2$$

Efficiency loss due to digital quantization:

$$\eta = \left(\frac{\sin(\pi/q)}{\pi/q} \right)^2$$

Right circularly polarized light:

$$E_{\text{in}} = \begin{bmatrix} Ex_{\text{in}} \\ Ey_{\text{in}} \end{bmatrix} = \begin{bmatrix} Ex_{\text{in}} \\ iEx_{\text{in}} \end{bmatrix}$$

Jones matrix representation of a half-wave plate:

$$E_{\text{out}} = \begin{bmatrix} \cos\beta & -\sin\beta \\ \sin\beta & \cos\beta \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 \\ 0 & e^{i\phi} \end{bmatrix} \cdot \begin{bmatrix} \cos\beta & \sin\beta \\ -\sin\beta & \cos\beta \end{bmatrix} \cdot \begin{bmatrix} Ex_{\text{in}} \\ iEx_{\text{in}} \end{bmatrix}$$

Equation Summary

Simplified half-wave plate:

$$E_{\text{out}} = \begin{bmatrix} Ex_{\text{in}}e^{i2\beta} \\ -iEx_{\text{in}}e^{i2\beta} \end{bmatrix}$$

Normalized Stokes parameter:

$$\eta_{m=0} = \cos^2\left(\frac{\pi\Delta nd}{\lambda}\right)$$

Potential steering orders:

$$\eta_{m=\pm 1} = \left[\frac{1 \mp S'_3}{2}\right] \sin^2\left(\frac{\pi\Delta nd}{\lambda}\right)$$

Lenslet steering angle:

$$\tan\theta = \frac{\Delta x}{f}$$

Standard sag formula:

$$OPD = (n - 1)d = \frac{\rho^2}{2f}$$

Maximum allowed lenslet steering angle:

$$OPD = \frac{\rho}{2} \tan(\theta_{\text{max}})$$

Mixed-lenslet steering angle:

$$\theta = \arctan\left(\frac{2\Delta x}{f}\right)$$

Focal spot diameter:

$$D = 2(f/\#)\lambda$$

Equation Summary

F-number:

$$f/\# = \frac{f}{D}$$

Numerical aperture:

$$NA = \frac{1}{2(f/\#)}$$

Convolution:

$$f * g = \int_{-\infty}^{\infty} f(\tau)g(t - \tau)d\tau$$

Digital convolution:

$$w(k) = \sum_j u(j)v(k - j + 1)$$

Fourier transform of a pupil plane field:

$$F\{E(x,y)\} = \iint E(x,y)e^{[-i2\pi(f_x x + f_y y)]} dx dy$$

Spatial heterodyne beat between the local oscillator and signal:

$$I = |u_t + u_{LO}|^2 = |u_t|^2 + |u_{LO}|^2 + u_t u_{LO}^* + u_t^* u_{LO}$$

Figure Sources

- p. 2 Chart derived from Google Ngram Viewer: LIDAR versus lidar versus LiDAR versus LADAR (1950–2008, English).
- p. 10 top http://en.wikipedia.org/wiki/Laser_safety. Public domain.
- p. 10 bottom http://en.wikipedia.org/wiki/Laser_safety. Figure created by Han-Kwang Nienhuys and released under license CC-BY-2.5.
- p. 11 http://en.wikipedia.org/wiki/Laser_safety. Figure created by Dahn and released under licence CC-BY-2.5.
- p. 18 Figure created by R. Pogge from ATRAN model data of Lord, S.D. (1992, NASA Technical Memor. 103957) from the Gemini Observatory website (<http://www.gemini.edu>).
- p. 19 top © Mike Willis December 26th, 2006.
- p. 19 bottom © The Worlds of David Darling.
- p. 27 Courtesy of Agilent Technologies.
- p. 30 http://en.wikipedia.org/wiki/Pink_noise. Figure created by Marrakkk and released under license CC BY-SA 3.0.
- p. 40 Courtesy of Matthew Dierking, Air Force Research Lab.
- p. 47 P. McManamon, “Review of ladar: A historic, yet emerging, sensor technology with rich phenomenology,” *Opt. Eng.* **51**(6), 060901 (2012) [doi: 10.1117/1.OE.51.6.060901].
- p. 52 National Research Council, *Laser Radar: Progress and Opportunities in Active Electro-Optical Sensing*, The National Academies Press, Washington, D.C., p. 63 (2014).
- p. 53 National Research Council *Laser Radar: Progress and Opportunities in Active Electro-optical Sensing*, The National Academies Press, Washington, D.C., p. 88 (2014).
- p. 54 R. C. Hardie, M. Vaidyanathan, and P. F. McManamon, “Spectral band selection and classifier design for a multispectral imaging laser radar,” *Opt. Eng.* **37**(3), 752–762 (1998) [doi: 10.1117/1.601907].

Figure Sources

- p. 58 Adapted from J. C. Marron and R. L. Kendrick, "Distributed aperture active imaging," *Proc. SPIE* **6550**, 65500A (2007) [doi: 10.1117/12.724769].
- p. 68 Courtesy of Jeff Hecht.
- p. 73 top <http://en.wikipedia.org/wiki/Mode-locking>.
Released under license CC BY-SA 3.0-MIGRATED.
- p. 79: R. V. Chimenti, M. P. Dierking, P. E. Powers, and J. W. Hausa, "Multiple chirp sparse frequency LFM ladar signals," *Proc. SPIE* **7323**, 73230N (2009) [doi: 10.1117/12.817865].
- p. 87 Courtesy of Intevac, Inc. <http://www.intevac.com/intevacphotonics/ebaps-technology-overview/>. Figure reprinted with permission.
- p. 93 Courtesy of How Stuff Works: <http://science.howstuffworks.com/gimbal1.htm>.
- p. 96 Courtesy of Nikon Small World: <http://www.microscopyu.com/articles/confocal/resonant-scanning.html>.
- p. 97 P. F. McManamon, P. J. Bos, M. J. Escuti, J. Heikenfeld, et al., "A review of phased array steering for narrow-band electro-optical systems," *Proc. IEEE* **97**(6), 1078–1096 (2009).
- p. 100 P. F. McManamon, P. J. Bos, M. J. Escuti, J. Heikenfeld, et al., "A review of phased array steering for narrow-band electro-optical systems," *Proc. IEEE* **97**(6), 1078–1096 (2009).
- p. 104 P. F. McManamon, P. J. Bos, M. J. Escuti, J. Heikenfeld, et al., "A review of phased array steering for narrow-band electro-optical systems," *Proc. IEEE* **97**(6), 1078–1096 (2009).
- p. 105 top S. R. Davis, G. Farca, S. D. Rommel, S. Johnson, and M. H. Anderson, "Liquid crystal waveguides: New devices enabled by >1000 waves of optical phase control," *Proc. SPIE* **7618**, 76180E (2010) [doi: 10.1117/12.851788].
- p. 105 bottom S. R. Davis, G. Farca, S. D. Rommel, S. Johnson, and M. H. Anderson, "Liquid crystal waveguides: New devices enabled by >1000 waves of optical phase control," *Proc. SPIE* **7618**, 76180E (2010) [doi: 10.1117/12.851788].

Figure Sources

- | | |
|---------------|--|
| p. 107 | P. F. McManamon, P. J. Bos, M. J. Escuti, J. Heikenfeld, et al., “A review of phased array steering for narrow-band electro-optical systems,” <i>Proc. IEEE</i> 97 (6), 1078–1096 (2009). |
| p. 108 top | P. F. McManamon, P. J. Bos, M. J. Escuti, J. Heikenfeld, et al., “A review of phased array steering for narrow-band electro-optical systems,” <i>Proc. IEEE</i> 97 (6), 1078–1096 (2009). |
| p. 108 bottom | P. F. McManamon, P. J. Bos, M. J. Escuti, J. Heikenfeld, et al., “A review of phased array steering for narrow-band electro-optical systems,” <i>Proc. IEEE</i> 97 (6), 1078–1096 (2009). |
| p. 109 | P. F. McManamon, P. J. Bos, M. J. Escuti, J. Heikenfeld, et al., “A review of phased array steering for narrow-band electro-optical systems,” <i>Proc. IEEE</i> 97 (6), 1078–1096 (2009). |
| p. 110 | E. A. Watson, “Analysis of beam steering with decentered microlens arrays,” <i>Opt. Eng.</i> 32 (11), 2665–2670 [doi: 10.1117/12.148100]. |
| p. 113 | Courtesy of Raytheon. |
| p.115 | http://en.wikipedia.org/wiki/Inertial_measurement_unit . Released under license CC BY-SA 3.0. |
| p. 118 | http://en.wikipedia.org/wiki/Aliasing . Released under license CC BY-SA 3.0. |
| p. 122 | P. McManamon, “Review of ladar: A historic, yet emerging, sensor technology with rich phenomenology,” <i>Opt. Eng.</i> 51 (6), 060901 (2012) [doi: 10.1117/1.OE.51.6.060901]. |

Bibliography

Alexander, S. B., *Optical Communication Receiver Design*, SPIE Press, Bellingham, Washington (1997) [doi: 10.1117/3.219402].

Beck, S., J. Buck, W. Buell, R. Dickinson, D. Kozlowski, N. Marechal, and T. Wright, "Synthetic aperture imaging laser radar: Laboratory demonstration and signal processing," *Appl. Opt.* **44**(35), 7621–7629 (2005).

Fox, C. S., Ed., *Active Electro-Optical Systems*, in *The Infrared & Electro-Optical Systems Handbook*, Vol. 6, J. S. Accetta and D. L. Shoemaker, Eds., Environmental Research Institute of Michigan, Ann Arbor, Michigan and SPIE Optical Engineering Press, Bellingham, Washington (1993).

S. B. Compagna, Ed., *Passive Electro-Optical Systems*, *The Infrared and Electro-Optical Systems Handbook*, Vol. 5, J. S. Accetta and D. L. Shoemaker, Eds., Environmental Research Institute of Michigan, Ann Arbor, Michigan and SPIE Optical Engineering Press, Bellingham, Washington (1993).

Fujii, T. and T. Fukuchi, Eds., *Laser Remote Sensing*, Taylor and Francis Group, Boca Raton, Florida (2005).

Henderson, S. W., P. Gatt, D. Rees, and M. Huffaker, "Wind LIDAR," Chapter 7 in *Laser Remote Sensing*, T. Fujii and T. Fukuchi, Eds., Taylor and Francis Group, Boca Raton, Florida (2005).

Jelalian, A. V., *Laser Radar Systems*, Artech House, Norwood, Massachusetts (1991).

Keyes, R. J., "Heterodyne and nonheterodyne laser transceivers," *Rev. Sci. Instrum.* **57**, 519 (1986).

Krause, B., J. Buck, C. Ryan, D. Hwang, P. Kondratko, A. Malm, A. Gleason, and S. Ashby, "Synthetic aperture lidar flight demonstration," CLEO 2011, Laser Applications to Photonic Applications, PDPB7, Optical Society of America, Washington, D.C. (2011).

Bibliography

Marron, J. C., R. L. Kendrick, N. Seldomridge, T. D. Grow, and T. A. Hoft, "Atmospheric turbulence correction using digital holographic detection: Experimental results," *Opt. Exp.* **17**, 11638 (2009).

McManamon, P. F. "Review of lidar: A historic, yet emerging, sensor technology with rich phenomenology," *Opt. Eng.* **51**(6), 060901 (2012) [doi: 10.1117/1.OE.51.8.089801].

McManamon, P. F., G. Kamerman, and M. Huffaker, "A history of laser radar in the United States," *Proc. SPIE* **7684**, 76840T (2010) [doi: 10.1117/12.862562].

National Research Council, *Laser Radar: Progress and Opportunities in Active Electro-Optical Sensing*, The National Academies Press, Washington, D.C. (2014).

Osche, G. R., *Optical Detection Theory for Laser Applications*, Wiley-Interscience, New York (2002).

Overbeck, J. A., M. S. Salisbury, M. B. Mark, and E. A. Watson, "Required energy for a laser radar system incorporating a fiber amplifier or an avalanche photodiode," *Appl. Opt.* **34**, 7724–7730 (1995).

Reibel, R. R., P. A. Roos, T. J. Berg, B. Kaylor, N. Greenfiend, Z. W. Barber, C. J. Renner, and W. R. Babbitt, "Ultra-broadband optical chirp linearization for precision length metrology applications," Optical Fiber Communication Conference, 2010, Optical Society of America, Washington, D.C., OThQ2 (2010).

Richmond, R. D. and S. C. Cain, *Direct-Detection LIDAR Systems*, SPIE Press, Bellingham, Washington (2010) [doi: 10.1117/3.836466].

Siegman, A. E., "The antenna properties of optical heterodyne receivers," *Proc. IEEE* **54**(10), 1350–1356 (1966).

Index

- 1D lidar, 3
- 1D slices, 41
- 1D waveguide, 66
- 1D waveguide lasers, 66
- 1/f noise, 7, 30
- 2D lidar, 2, 42
- 3D lidar, 3
- 3D mapping, 4, 43
- 3D point cloud images, 127
- 3D scanning lidars, 43
- 3D shape, 127
- 3D shape information, 43
- 3D voxels, 122, 127

- absolute motion, 115
- absorption, 109
- accelerometers, 115
- acousto-optic devices, 85
- acousto-optic modulator, 85
- active cladding layer, 105
- active multispectral imaging, 54
- adaptive optics mirror, 22
- aero-optical effect, 21
- aero-optical layer, 21
- afterpulsing, 32
- Airy disk, 92
- all-positive lenslet approach, 112
- amplification stages, 75
- amplitude variations, 22
- angle information, 41
- angular distribution, 41
- angular resolution, 59
- angular rotation, 50
- applied voltage, 99
- archeological mapping, 4
- artificial gain medium, 68
- atmospheric absorption, 18
- atmospheric coherence time, 20
- atmospheric loss, 19
- atomic constituents, 52
- autocorrelation, 124
- avalanche gain, 31
- avalanche photodiode, 30–31, 44–45, 82

- background noise, 29
- band structure engineering, 68
- bandwidth of the lidar, 40
- beam, 8
- beam quality, 74
- birefringence, 99
- bistatic lidar, 1, 12, 61
- blue shift, 90
- bulk defects, 30
- bulk solid state lasers, 64

- carrier-to-noise ratio, 37
- cascade diode lasers, 67
- cellular network planning, 4
- central limit theorem, 26
- chip, 80, 102
- chirping, 47
- circularly polarized gratings, 107
- cladding, 65
- Class 1 lasers, 11
- Class 2 lasers, 11
- Class 3 lasers, 11
- Class 3B lasers, 11
- Class 4 lasers, 11
- clock frequency drifts, 117

Index

- CO₂, 6, 18
- CO₂ laser, 61
- code division multiple
 - access, 80
- coherence converters, 61
- coherence length, 9
- coherent detection, 84
- coherent detection system, 41
- coherent lidar, 1, 3, 7, 24, 28
- coherent receiver, 34
- coincidence processing, 45
- coincident returns, 32
- conformal steering
 - approach, 95
- constant fraction
 - detection, 119
- controlled laser modes, 66
- convolution, 117
- corner cubes, 15
- cross section, 14
- cross section of a corner cube, 16

- dark current, 30
- dark current noise, 30
- dark frame subtraction, 30
- dead time, 83
- deflection efficiency, 85
- deformable mirror, 22
- degree of polarization, 55
- delta function, 89
- detector angular subtense, 3, 92, 114
- differential absorption
 - lidar, 5
- diffraction limit, 8

- digital beam steering, 109
- digital convolution, 117
- digital holography, 7, 33
- dihedral, 16
- diode, 65
- diode lasers, 61, 67
- direct detection, 7, 23, 25, 31
- direct detection lidar, 24, 81
- direct detection receiver, 34
- direction of the
 - illumination, 55
- discrete Poisson
 - probability density function, 26
- Doppler lidar, 5
- Doppler shift, 48, 77, 90, 126
- Doppler shift
 - measurement time, 90
- Doppler spectrum, 41
- doublet pulse waveform, 78
- dual-frequency liquid crystals, 99

- edge-emitting diode lasers, 67
- effective aperture, 59
- effective grayscale, 83
- effective range profile, 83
- effectively apodized
 - aperture array, 59
- electrical singal-to-noise ratio, 25
- electro-optical scanning, 106

Index

- electron-bombarded
 - active-pixel sensor, 87
- electron-bombarded semiconductor, 87
- electronic levels, 53
- electrowetting, 104
- emission spectrum, 52
- emission wavelengths, 52
- end pumping, 69
- energy received equation, 23
- Er:YAG lasers, 64
- excess noise, 30
- excess noise factor, 30
- eye damage threshold, 10
- eye safety, 10
- eyesafe region, 64

- fast Fourier transform, 121
- fast-steering mirror, 94
- ferroelectric liquid
 - crystals, 109
- fiber laser, 65
- fiber lasers, 61
- fiber tip damage, 66
- field of view, 3
- fill factor, 103
- fixed-pattern noise, 30
- flash image, 115
- flash imagers, 44
- flash imaging, 1
- flash imaging with lidar, 88
- flash lidar, 12
- flashlamps, 69
- flicker noise, 30
- flyback region, 100
- FM chirp, 77

- focal spot, 116
- focal spot diameter, 114
- focusing ability, 114
- forestry lidar, 4
- forward-looking infrared cameras, 86
- four-level laser, 63
- Fourier transform, 121
- frequency chirping, 89
- frequency division
 - multiple access, 80
- frequency line shape, 73
- frequency shifting, 76
- Fresnel loss, 109
- Fried parameter, 20
- fringing fields, 100–101
- frozen atmosphere
 - assumption, 20–21

- gain medium, 75
- gain region, 65
- gamma distribution, 26
- gated 2D camera, 42
- gated 2D imaging, 87
- gating circuitry, 42
- Gaussian beam, 8
- Gaussian illumination, 98
- Gaussian probability
 - density function, 26
- Gaussian-shaped pulse, 118
- Geiger-mode APD, 23, 83, 115
- Geiger-mode APD camera, 122
- Geiger-mode lidar, 125
- general image quality
 - equation, 123
- geometrical optics, 114

Index

- gimbal pointing, 93
- gimbals, 21, 93
- GMAPD detection events, 32
- GMAPD flash imagers, 45
- GMAPD-based flash imaging, 45
- gyroscopes, 115
- half-power point, 98
- hard targets, 54
- heterodyne detection, 23, 89
- heterodyne detection receiver, 34
- heterodyne mixing efficiency, 35
- HgCdTe, 46
- high range resolution, 47
- high time–bandwidth product waveform, 78
- high time–bandwidth product waveforms, 77
- high-duty-cycle waveforms, 67
- high-gain laser media, 75
- high-index oil, 104
- hill climbing algorithm, 58
- HoTm:YAG lasers, 64
- illuminator aperture, 12
- imposing the small motion, 116
- in-plane steering, 105
- index of refraction, 20
- index of refraction change, 72
- inertial measurement unit, 115
- interband lasers, 67
- interference, 17, 56
- intermediate frequency, 36–37
- inverse synthetic aperture lidar, 50
- inverse synthetic aperture radar, 50
- irradiation, 52
- isolation, 75
- isoplanatic angle, 20
- isoplanatic patch, 20
- Johnson noise, 27
- Johnson’s criteria, 123
- Kerr-effect EO crystals, 106
- KTN, 106
- LADAR, 2
- Lambertian scattering, 29
- large-angle steering, 113
- largest angle that can be steered, 98
- laser beam divergence, 8
- laser designators, 6
- laser diodes, 6
- laser imaging and ranging, 72
- laser linewidth, 74
- laser pulse, 77
- laser radar, 2
- laser radiation, 10
- laser range finders, 6
- laser remote sensing, 2
- laser resonator, 62

Index

- laser signal, 1
- laser vibrometer, 48
- laser vibrometer signal
 - processing, 126
- laser vibrometry, 126
- laser waveform, 1
- laser-induced breakdown
 - spectroscopy, 5, 52
- laser-induced
 - fluorescence, 5
- laser-induced fluorescence
 - lidar, 53
- leading edge detection, 119
- lensless imaging, 121
- lenslet-based beam
 - steering, 110
- Level 0 (L0), 125
- Level 1 (L1), 125
- Level 2 (L2), 125
- Level 3 (L3), 125
- Level 4 (L4), 125
- Level 5 (L5), 125
- lidar, 1–2
- lidar clock frequency, 117
- lidar cross section, 15
- lidar data processing
 - stages, 125
- lidar development, 6
- lidar image stabilization, 88
- lidar peak power, 66
- lidar pulse, 89
- lidar range equations, 14
- lidar range resolution, 89
- lidar waveform, 77
- lidar wavelengths, 18
- LIMAR, 86
- LiNbO₃, 106
- line broadening, 73
- linear EO crystals, 106
- linear EO effect, 106
- linear FM waveforms, 80
- linear frequency
 - modulation, 79, 89
- linear-mode APD, 23, 46, 82, 122
- liquid crystal layer, 100
- liquid crystal polarization
 - gratings, 108
- liquid crystal steering
 - device, 99
- local oscillator, 28, 33, 37
- local oscillator power, 24
- long-focal-length optics, 114
- long-integration-time
 - framing cameras, 86
- low energy per pulse, 83
- low-index water, 104
- master oscillator, 75
- master oscillator, power
 - amplifier, 75
- matched filter, 117
- maximum permissible
 - exposure, 10–11
- measured range profile, 117
- MEMS mirror, 94
- MEMS mirror arrays, 103
- MEMS-mirror-based
 - steering, 103
- micro-actuators, 103
- microlens array, 110
- micromotion, 116
- Mie scattering, 19
- mixed set of lenslet arrays, 112

Index

- mixed-lenslet steering, 112
- mixed-pixel effect, 55
- mode-locked lasers, 73
- modulo 2π beam steering, 97
- modulo 2π phase profile, 97
- monochromaticity, 9
- monostatic lidars, 1, 12
- Mueller matrix
 - characterization, 55
- multiple input, multiple output lidars, 12
- multiple pulses, 83
- multiple subapertures, 124
- multiple tagged laser transmitters, 76
- multiple-aperture lidars, 57
- multiple-input, multiple-output lidars, 57, 59
- nanofabricated phased arrays, 102
- National Imagery Interpretability Rating Scale, 123
- Nd:YAG designators, 64
- Nd:YAG lasers, 61, 64
- negative binominal distribution, 26
- nitrogen, 18
- noise current, 24
- noise equivalent photons, 87
- noise equivalent temperature difference, 87
- noise probability density functions, 26
- nonlinear effects, 70
- number of photons, 32
- numerical aperture, 114
- Nyquist criterion, 118
- Nyquist noise, 27
- Nyquist sampling theorem, 118
- object identification, 127
- oil and gas exploration, 4
- optical breakdown, 52
- optical parametric amplifier, 70
- optical parametric oscillator, 70
- optical path difference, 97
- optical phased arrays, 57
- optical power received, 14
- optimal wavelengths, 54
- order of the blaze, 104
- out-of-plane steering, 105
- overall laser efficiency, 75
- oxygen, 18
- pattern of changes, 91
- peak detection, 119
- peak-power limitation, 65
- periodically poled lithium niobate, 70
- phase center, 49
- phase change, 22
- phase correction, 22
- phase corrector plates, 110
- phase delay, 111
- phase function, 92
- phase irregularities, 56
- phase profile, 92

Index

- phase screen, 92
- phase shift interferometry, 7
- phase steps, 101
- phase up the multiple subapertures, 60
- phased array of phased-array imaging, 57
- photon, 28
- photonic crystal fiber lasers, 66
- photonic crystal fibers, 65–66
- pink noise, 30
- pitch, 115
- Pockels cell, 47, 71–72
- point spread function, 3, 92
- polarimetric measurements, 55
- polarization, 13, 55
- polarization-based flash lidar imaging, 47
- polarization-encoded imaging, 44
- polarizing beamsplitter, 13
- polygon scanning, 96
- polymers, 99
- polypulse waveform, 78
- population inversion, 63
- power received equation, 23
- precise object identification, 4
- preform, 65
- pseudo-random-coded modulation, 77, 89
- pseudo-random-coded waveform, 80
- pulse doublet, 77–78
- pulse position modulation, 80
- pulsed ultraviolet sources, 53
- Q, 71
- Q switching, 71
- Q-switched lasers, 42
- quadratic Kerr effect, 106
- quadrature detection, 36
- quantization loss, 101
- quantum cascade lasers, 68
- quantum defect, 63
- quarter-wave plate, half-wave plate, quarter-wave plate, 107
- quasi-cw operation, 64
- quasi-four-level lasers, 63
- radiance, 14
- radiance per wavelength, 29
- Raman lidar, 5
- random interference, 9
- range, 3
- range accuracy, 120
- range depth, 79
- range Doppler imaging, 51
- range information, 41, 47
- range measurement accuracy, 117
- range precision, 120
- range profile, 40
- range resolution, 3, 40, 120
- range uncertainty, 79
- range-gated active imaging, 42

Index

- range-only lidar, 3
- range-resolved data, 53
- range/species resolution, 53
- Rayleigh scattering, 19
- read-out integrated circuit, 82
- receive aperture, 12
- receive-only arrays, 57
- receiver, 13
- reconstruct the range profile, 118
- red shift, 90
- reference beam, 36
- reference frame adjustment, 50
- reflective or transmissive modes, 104
- relative motion, 115
- relative time delay, 40
- required transmitted energy, 23
- retina, 10
- RF modulation, 81
- Risley gratings, 95
- Risley prisms, 95
- roll, 115
- rotating phase, 108
- rotating polygons, 96
- rotation speed, 41
- route planning, 4

- sag formula, 111
- sampling density, 116
- saturable absorber, 71
- sawtooth phase profile, 97
- scanning lidars, 122
- scatter, 109
- seeding a laser, 74

- series of sine waves, 118
- shot noise, 27–28
- side pumping, 64, 69
- signal current, 24
- signal-to-noise equation, 25, 39
- signal-to-noise ratio, 23–24, 37, 82
- single detector, 51
- slip rings, 93
- solid angle return, 16
- solid state lasers, 61
- spatial coherence, 8
- spatial fringes, 35
- spatial heterodyne detection, 7, 33, 38–39, 86, 124
- spatial heterodyne imaging, 58
- speckle, 9, 17
- speckle imaging lidar, 56
- speckle pattern, 56
- speckle realization, 17
- spectroscopic analysis, 52
- spontaneous emission, 63
- spurious sidelobes, 102
- stabilization, 43
- stable resonators, 62
- standard framing cameras, 47
- steering angle, 110
- steering efficiency, 101
- steering efficiency versus angle, 100
- steering loss, 109
- steering prisms, 95
- steering to discrete angles, 110
- stimulated emission, 63

Index

- stitching, 115
- stochastic parallel
 - gradient descent, 58
- Stokes vector, 55
- stretch processing, 79
- subapertures, 59
- sum and difference
 - frequency generation, 70
- super-Gaussian, 8
- super-Gaussian beam, 8
- surface of the eye, 10
- surface roughness, 55
- surveying, 4
- switching time, 99
- synthetic aperture
 - imagers, 49
- synthetic aperture lidar, 3, 49
- synthetic aperture radar, 49

- target reflectivity, 14
- temporal coherence, 9
- temporal heterodyne
 - detection, 7, 33–35, 84–85
- temporal heterodyne imaging, 58
- temporal heterodyning, 48
- ternary steering, 108
- thermal noise, 27
- three-level laser, 63
- threshold detectors, 119
- Ti:sapphire laser, 73
- tilting, 103
- time analysis, 53
- time division multiple
 - access, 80
- time gating, 52
- time tag, 89
- time–bandwidth product, 78
- Tm:YAG lasers, 64
- tomographic images, 41
- transmissive beam
 - steering, 104
- transmit and receive
 - arrays, 57
- transmit aperture, 12
- transmit/receive isolation, 13
- transmitter, 13
- transmitter array, 76
- triangular, 16
- turbulence, 21

- unambiguous range, 47, 91
- uniform illumination, 98
- unstable resonators, 62

- velocity, 90
- vertical-cavity surface-emitting lasers, 67
- vibration, 48
- vibrational image, 48
- virtual phase screen, 22
- volume holographic gratings, 113

- water vapor, 18
- waveform generator, 1
- waveguide coupler, 105
- wide-angle steering, 110

- yaw, 115

- zero crossing, 95



Paul McManamon owns Exciting Technology LLC and is the Technical Director of Ladar and Optical Communications Institute (LOCI) at the University of Dayton. Dr. McManamon recently chaired the National Academy of Sciences (NAS) study “Laser Radar: Progress and Opportunities in Active Electro-Optical Sensing,” released in March, 2014. Dr.

McManamon co-chaired the NAS study, “Optics and Photonics: Essential Technologies for Our Nation,” released in August, 2012. This study recommended the National Photonics Initiative. He also served as Vice Chair to the 2010 NAS study “Seeing Photons.”

Dr. McManamon started at Wright-Patterson Air Force Base in May 1968 and retired from the same in 2008 as Chief Scientist of the Sensors Directorate, U.S. Air Force Research Laboratory (AFRL). Previous to that he was Senior Scientist of EO Sensors, and before that, acting chief scientist for Avionics, also at AFRL.

In 2006 he received the Meritorious Presidential Rank Award. He has participated in three Air Force Scientific Advisory Board studies. He was instrumental in the development of laser flash imaging to enhance EO target recognition range by a factor of 4 or 5. Dr. McManamon was the 2006 President of SPIE. He served on the SPIE Board of Directors for seven years and on the SPIE Executive Committee for four years. In 1998, Dr. McManamon and his co-authors received the IEEE W.R.G. Baker award for best paper in *any* refereed IEEE journal or publication. Dr. McManamon is a Fellow of SPIE, IEEE, OSA, AFRL, MSS, AIAA, and DEPS.