

SPIE PRESS | Field Guide



**SPIE**

Field Guide to

# **Terahertz Sources, Detectors, and Optics**

***Créidhe O'Sullivan  
J. Anthony Murphy***

**SPIE Terms of Use:** This SPIE eBook is DRM-free for your convenience. You may install this eBook on any device you own, but not post it publicly or transmit it to others. SPIE eBooks are for personal use only. For details, see the SPIE [Terms of Use](#). To order a print version, [visit SPIE](#).

**SPIE.**

Field Guide to

# **Terahertz Sources, Detectors, and Optics**

Créidhe O'Sullivan  
J. Anthony Murphy

SPIE Field Guides  
Volume FG28

John E. Greivenkamp, Series Editor

**SPIE**  
**PRESS**

Bellingham, Washington USA

Library of Congress Cataloging-in-Publication Data

O'Sullivan, Créidhe M. M.

Field guide to terahertz sources, detectors, and optics /  
Créidhe O'Sullivan, J. Anthony Murphy.

p. cm. – (The field guide series ; FG28)

Includes bibliographical references and index.

ISBN 978-0-8194-9167-1

1. Millimeter wave devices—Handbooks, manuals, etc.
2. Terahertz technology—Handbooks, manuals, etc.
3. Infrared equipment—Handbooks, manuals, etc.
4. Submillimeter waves—Handbooks, manuals, etc.

I. Murphy, J. Anthony. II. Title.

TK7876.5.O88 2012

621.381'33—dc23

2012012372

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](mailto:books@spie.org)

Web: <http://spie.org>

Copyright © 2012 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 easily 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](mailto: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 all of the titles in the Field Guide Series:

### *Field Guide to*

*Adaptive Optics*, Robert Tyson & Benjamin Frazier

*Atmospheric Optics*, Larry Andrews

*Binoculars and Scopes*, Paul Yoder, Jr. &  
Daniel Vukobratovich

*Diffraction Optics*, Yakov Soskind

*Geometrical Optics*, John Greivenkamp

*Illumination*, Angelo Arecchi, Tahar Messadi, &  
John Koshel

*Image Processing*, Khan M. Iftakharuddin &  
Abdul Awwal

*Infrared Systems, Detectors, and FPAs, Second Edition*,  
Arnold Daniels

*Interferometric Optical Testing*, Eric Goodwin &  
Jim Wyant

*Laser Pulse Generation*, Rüdiger Paschotta

*Lasers*, Rüdiger Paschotta

*Microscopy*, Tomasz Tkaczyk

*Optical Fabrication*, Ray Williamson

*Optical Fiber Technology*, Rüdiger Paschotta

*Optical Lithography*, Chris Mack

*Optical Thin Films*, Ronald Willey

*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

*Visual and Ophthalmic Optics*, Jim Schwiegerling

---

### *Field Guide to Terahertz Sources, Detectors, and Optics*

## Field Guide to THz

---

The region of the electromagnetic spectrum between microwaves and infrared radiation has come to be known as the “THz gap,” mainly due to the lack of readily available laboratory sources and detectors. For many years technology development was driven by astronomers and planetary scientists, but other potential uses, particularly in medical and security applications, have led to increased activity by the mainstream physics and engineering community in recent times. Because diffraction is important at these frequencies, THz systems cannot be successfully designed using traditional optical techniques alone.

The primary objective of this Field Guide is to provide the reader with a concise description of the quasi-optical techniques used at THz frequencies, as well as the basic principles of operation of the most common THz system components in use today. More detailed accounts of specific devices can be found in the bibliography and references therein.

We would like to thank our families and our colleagues at NUI Maynooth, in particular Neil Trappe, Marcin L. Gradziel, and Ian McAuley of the THz Optics group and also Stafford Withington of the Cavendish Laboratory at Cambridge.

Créidhe O’Sullivan  
J. Anthony Murphy  
Department of Experimental Physics  
National University of Ireland,  
Maynooth





## Table of Contents

<b>Glossary of Symbols and Acronyms</b>	<b>xi</b>
<b>Introduction</b>	<b>1</b>
The THz Band	1
The THz Gap	2
THz Absorption in Air	3
Reflection and Transmission	4
<b>THz Sources</b>	<b>5</b>
Natural Sources of THz Radiation	5
THz Generation Techniques	6
Gunn Diodes	8
IMPATT Diodes	9
TUNNETT Diodes	10
Resonant Tunnel Diodes	11
Difference Frequency Generation	12
Electro-optic Crystals (Optical Rectification)	13
Optical Parametric Oscillators	15
Frequency Multipliers	16
Photoconductive Antennas	17
Photomixing	18
Optically Pumped Far-IR Gas Lasers	19
p-Type Germanium Lasers	20
Quantum Cascade Lasers	21
Gyrotrons	22
Synchrotrons	23
Free-Electron Lasers	24
Backward Wave Oscillators	25
Smith–Purcell Emitters	26
<b>THz Detectors</b>	<b>27</b>
THz Detection Techniques	27
Responsivity and Signal-to-Noise Ratio	29
Noise Equivalent Power	30
Shot and Thermal Noise	31
Extrinsic Semiconductor Detectors	32
Photoconductive Detectors	33
Photomixers	34
Schottky Diodes	35
Schottky Diode Mixers	37
SIS Mixers	38
Electro-optic Sampling	39
Semiconductor Bolometers	41
Microbolometer Arrays	43

## Table of Contents

Transition-Edge Sensors	44
SQUIDs	45
Hot-Electron Bolometers	47
Hot-Electron Bolometer Mixers	48
Pyroelectric Detectors	49
Golay Cells	50
<b>THz Optics</b>	<b>51</b>
Gaussian Beam Propagation	51
Complex Radius of Curvature	52
Gaussian Beam Parameters	53
Beamwidth	55
Edge Taper	56
Truncation and Spillover (Gaussian Beams)	57
Truncation and Spillover (Non-Gaussian Beams)	58
Confocal Distance	59
Far-Field Divergence	60
Ray Matrices and Gaussian Beam Transformation	61
Higher-Order Modes (Cylindrical Coordinates)	64
Higher-Order Modes (Rectangular Coordinates)	65
Mode Coefficients	66
Power Coupling Efficiency	69
Gaussianity	70
Mismatched Beams and Tolerancing	71
Scattering Matrix Formulism	72
Transmission Matrices	73
Linear Scattering Operators	74
Truncation at an Aperture	75
Perfect Lenses and Pure Propagation	76
Reflections at Dielectric Interfaces (Lens Surfaces)	77
Standing Waves in Horn-Fed Systems	78
Cascading Scattering and Transmission Matrices	79
Off-Axis Mirrors (Ellipsoidal)	80
Off-Axis Mirrors (Parabolic)	81
Off-Axis Mirrors (Distortion and Cross-Polarization)	82
Polarizing Grids	83
Roof Mirrors as Polarization Rotators	84

## Table of Contents

---

Dual-Beam Interferometers (Tunable Filters)	85
Diffraction Losses in Interferometers	86
Diplexers and Multiplexers	87
Four-Port Dual-Beam Interferometer (Diplexer)	88
Horn Antenna Feeds	89
Corrugated Conical Horns (Scalar Feed)	90
Smooth-Walled Horns (Pyramidal and Diagonal)	91
Conical Smooth-Walled Horns (Single and Dual Mode)	92
Shaped Horns and Multimode Feeds	93
Lens Antennas	94
System Design	95
Modeling Techniques	96
<b>THz Applications</b>	<b>97</b>
THz Imaging	97
THz Spectroscopy	99
THz Time-Domain Spectroscopy	100
<b>Equation Summary</b>	<b>101</b>
<b>Bibliography</b>	<b>114</b>
<b>Index</b>	<b>117</b>



## Glossary of Symbols and Acronyms

---

$A$	absorbance
$A$	detector area
$A$	profiled horn parameter
$a$	aperture radius
$a$	horn aperture side length
$a$	semi-major axis of an ellipse
$a$	wire radius
$A_n, A_n^h, A_n^v$	$n^{\text{th}}$ mode coefficient of beam $a$ , $n^{\text{th}}$ mode coefficient of orthogonal components
$\mathbf{a}_n$	mode coefficients for beam $n$ at the input port
ac	alternating current
$\mathbf{A}_{12}, \mathbf{A}_{21}$	scattering matrices for an absorbing stop
$\mathbf{B}$	magnetic field
$B$	magnetic field strength
$b$	horn aperture side width
$b$	semi-minor axis of an ellipse
$B_n, B_n^h, B_n^v$	$n^{\text{th}}$ mode coefficient of beam $b$ , $n^{\text{th}}$ mode coefficient of orthogonal components
$\mathbf{b}_n$	mode coefficients for beam $n$ at the output port
BWO	backward wave oscillator
$C$	capacitance
$C$	heat capacity
$c$	speed of light
$C_e, C_{ph}$	electron heat capacity, phonon heat capacity
c.c.	complex conjugate
CMOS	complementary metal-oxide-semiconductor
CW	continuous wave
$D$	detectivity
$d$	diameter
$d$	distance between the focal point and the apex of a lens antenna
$d$	propagation distance

## Glossary of Symbols and Acronyms

---

$D^*$	specific detectivity
$\mathbf{D}_{12}, \mathbf{D}_{21}$	scattering matrices for pure propagation
dc	direct current
DFG	difference frequency generation
$\mathbf{E}, \mathbf{E}_a$	electric field
$\mathbf{E}(t), \mathbf{E}(r)$	electric field
$E, E_0$	electric field amplitude
$E(x, y, z), E(\mathbf{r})$	electric field amplitude
$e$	elementary charge
$E_a, E_b$	electric field amplitude
$\mathbf{E}_{ap}$	electric field across an aperture
$\mathbf{e}_F$	vector field
$\mathbf{e}_G$	vector Gaussian field
$E_g, E_C, E_V, E_F$	energy gap, conduction band energy, valence band energy, Fermi energy
$\hat{\mathbf{e}}_h, \hat{\mathbf{e}}_v$	unit orthogonal vectors
$E_i, E_j, E_k$	electric field amplitude
$E_{inc}, E_{refl}, E_{trans}$	incident, reflected, transmitted electric field
$E_L, E_H$	light-, heavy-hole energy
$E_{OP}$	optical phonon energy
$E_{out}$	output electric field
$E_{p,t}, E_{p,r}, E_{p,i}$	transmitted, reflected, incident electric field amplitude for p-polarization
$E_{s,t}, E_{s,r}, E_{s,i}$	transmitted, reflected, incident electric field amplitude for s-polarization
$E_{THz}$	terahertz electric field amplitude
$E_p^+, E_p^-$	forward and reverse traveling $p^{\text{th}}$ symmetric Laguerre mode
EO	electro-optic
$F$	Fresnel number
$f$	focal length
FEL	free-electron laser
FS	free space
FWHM	full width half maximum

## Glossary of Symbols and Acronyms

---

$G$	thermal conductance
$g$	wire spacing
GBMA	Gaussian beam mode analysis
GBT	Gaussian beam telescope
$\mathbf{H}, \mathbf{H}_a, \mathbf{H}_b$	magnetic field
$h$	Planck's constant
$H_m(x)$	$m^{\text{th}}$ Hermite polynomial
HEB	hot electron bolometer
HIFI	Heterodyne instrument for the far-infrared
HPBW	half-power beamwidth
$I$	current
$I$	intensity, transmitted intensity
$\mathbf{I}$	identity matrix
$i$	$\sqrt{-1}$
$i$	angle of incidence
$I_c$	critical current
$I_x, I_y$	$x, y$ component of intensity
$I_0$	incident intensity
IF	intermediate frequency
IMPATT	impact ionization avalanche transit time
IR	infrared
$J$	free current density
$J$	rotational state number
$J$	screening current
$K$	power coupling coefficient
$k$	wavenumber ( $2\pi/\lambda$ )
$K_{ab}$	power coupling between two beams
$k_B$	Boltzmann constant
$K_G$	power coupling to a pure Gaussian (Gaussicity)
$K_{lost}$	power lost due to truncation
$\mathbf{k}_P, \mathbf{k}_S, \mathbf{k}_I$	pump wavevector, signal wavevector, idler wavevector
$K_{spillover}$	power lost due to spillover
$K_{Xpol}$	power scattered into the cross-polar direction

## Glossary of Symbols and Acronyms

---

$L$	inductance
$L, L_x, L_y$	slant length, slant length along $x$ direction, slant length along $y$ direction
$l$	grating period
$L_{mn}$	operator representing a transfer function
$\mathbf{L}_{nn}$	coupling to reflected modes at port $n$
$L_p^m$	generalized Laguerre polynomials
$l_w$	walk-off length
$l_c$	coherence length
$l_1, l_2$	path length
LO	local oscillator
LT-GaAs	low-temperature-grown gallium arsenide
$M$	magnification
$\mathbf{M}_n$	$n^{\text{th}}$ matrix
$m_e$	electron mass
$N$	multiplication factor of a frequency multiplier
$n$	number density
$n, n_{\text{signal}}, n_{\text{noise}}$	number of photons, signal photons, noise photons detected
$n, n_1, n_2$	refractive index
$n$	rms noise level
$n$	spectral order
$n_{\text{optical}}, n_{\text{THz}}$	optical, terahertz refractive index
NDR	negative differential resistance
NEP	noise equivalent power
NTD	neutron-transmutation doped
OP	optical phonon
OPO	optical parametric oscillator
OPTL	optically pumped THz laser
$P, P_i, P_0, \mathbf{P}$	polarization
$P(r)$	power contained within a radius $r$
$P$	vertex-focus distance of a parabola
$P_{\text{noise}}, P_{\text{signal}}$	noise power, signal power
$P_{\infty}$	total power



## Glossary of Symbols and Acronyms

---

PCA	photoconductive antenna
p-Ge	p-type germanium laser
$q$	charge
$q(z), q_0$	complex radius of curvature (also known as complex beam parameter or Gaussian beam parameter)
$q_{in}, q_{out}$	complex radius of curvature of input, output beam
$q_1, q_2$	complex radius of curvature of beam 1, 2
QCL	quantum cascade laser
$R(z)$	Gaussian beam phase radius of curvature
$R$	radius
$R$	reflectivity
$R$	resistance
$R$	surface radius of curvature
$r$	complex reflection amplitude coefficient
$R_{in}, R_{out}$	phase radius of curvature
$r_{in}, r_{out}$	ray displacement
$R_{surf}$	surface radius of curvature
$r_T$	truncation radius
$R_V$	responsivity
$R_1, R_2$	surface radius of curvature
$R_1, R_2$	distances to foci of an ellipse
RF	radio frequency
rms	root-mean-square
RTD	resonant tunnel diode
$\mathbf{S}, \mathbf{S}^a, \mathbf{S}^b, \mathbf{S}^c$	full scattering matrix
$s$	phase error
$\mathbf{S}_{11}, \mathbf{S}_{12}, \mathbf{S}_{21}, \mathbf{S}_{22}$	scattering matrices
SIS	superconductor–insulator–superconductor
SNR	signal-to-noise ratio
$\text{SNR}_{dB}$	signal-to-noise ratio in dBs

## Glossary of Symbols and Acronyms

---

S/N	signal-to-noise ratio
SPR	Smith–Purcell radiation
SQUID	superconducting quantum interference device
$T$	temperature
$T$	transmittance
$T$	fractional power transmission
$t$	complex transmission amplitude coefficient
$\mathbf{T}, \mathbf{T}^a, \mathbf{T}^b, \mathbf{T}^c$	full transmission matrix
$T_b, T_s$	bath, heat sink temperature
$T_c$	critical temperature
$T_e$	edge taper
$T_e$ (dB)	edge taper expressed in dB
$T_e, T_{ph}$	electron temperature, phonon temperature
$\mathbf{T}_{11}, \mathbf{T}_{12}, \mathbf{T}_{21}, \mathbf{T}_{22}$	transmission matrices
TCR	thermal coefficient of resistance
TE, TM	transverse electric, magnetic
TES	transition edge sensor
THz-TDS	terahertz time-domain spectroscopy
TUNNETT	tunnel injection transit time
$u(x, y, z), u_0$	non-plane-wave part of the electric field, constant
$V, V_{rms}$	voltage, root-mean-square voltage
$v$	vibrational state number
$V_{bi}$	built-in potential
$V_{gap}$	gap voltage
$w, w(z)$	Gaussian beam radius
$w_m$	beam radius at a mirror
$w_0, w_{0,a}, w_{0,b}$	Gaussian beam waist radius
$w_{0,in}, w_{0,out}$	Gaussian beam waist radius
WG	waveguide
$x$	path length
$x_c$	point of intersection, $x$ coordinate

## Glossary of Symbols and Acronyms

---

$y_c$	point of intersection, $y$ coordinate
$z_c$	confocal distance, Rayleigh range
$\alpha$	absorption coefficient
$\alpha$	diffraction parameter
$\alpha$	incident beam polarization angle
$\alpha, \alpha(T)$	thermal coefficient of resistance
$\alpha$	tilt of grid wires
$\beta$	mode balance constant
$\beta_{trans}, \beta_{refl}$	transmitted, reflected beam polarization direction
$\chi$	electron affinity (volts)
$\chi^{(2)}, \chi_{ijk}^{(2)}$	second-order susceptibility
$\Delta$	path length difference
$\Delta_{IF}$	IF bandwidth
$\Delta_{max}$	path length difference for maximum transmission
$\Delta_{min}$	path length difference for minimum transmission
$\Delta_{SSB}$	single-sideband path length difference
$\Delta_1, \Delta_2$	energy gap of a semiconductor
$\Delta E$	energy gap
$\Delta x$	lateral shift
$\Delta z$	relative displacement
$\Delta z$	position of beam waist behind aperture
$\Delta \theta$	cone angle
$\Delta \nu$	bandwidth
$\delta \nu$	frequency of successive transmission bands
$\epsilon_0$	permittivity of free space
$\Phi_0$	flux quantum
$\phi$	grid inclination angle
$\phi_c$	critical emission angle
$\phi_m, \phi_{sc}, \phi_B$	metal, semiconductor, barrier work function
$\phi_0(z), \phi_{in}, \phi_{out}$	Gaussian beam phase slippage
$\gamma$	Lorentz factor ( $1/\sqrt{1 - \beta^2}$ ), $\beta = v/c$ , $v$ is velocity

## Glossary of Symbols and Acronyms

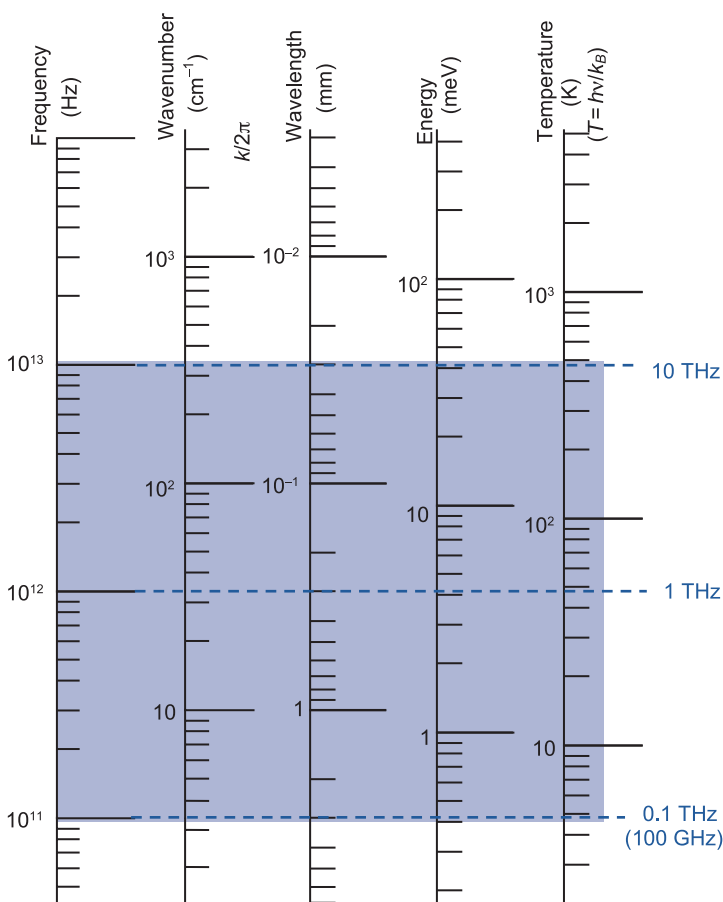
---

$\gamma$	reflection coefficient at an interface
$\eta_{\text{spillover}}$	Gaussian beam spillover efficiency
$\lambda$	wavelength
$\lambda_{\text{FEL}}$	wavelength of radiation from an FEL
$\lambda_{\text{IF}}, \lambda_{\text{LO}}, \lambda_s$	IF, LO, signal wavelength
$\lambda_w$	wiggler magnet spacing wavelength
$\nu$	frequency
$\nu_{\text{IF}}, \nu_{\text{LO}}, \nu_s$	IF, LO, signal frequency
$\nu_{\text{THz}}, \nu_{\text{pump}}$	terahertz, pump frequency
$\theta$	angle from normal
$\theta$	angle with respect to propagation axis
$\theta$	incident field polarization angle
$\theta$	grid tilt angle
$\theta'$	projected grid tilt angle
$\theta_c$	critical angle
$\theta_{\text{in}}, \theta_{\text{out}}$	ray angle with respect to normal
$\theta_0$	asymptotic beam growth angle
$\theta_1, \theta_2$	angle of incidence, reflection
$\sigma$	absorption cross-section
$\tau$	transmission
$\tau_{\text{diff}}$	diffusion time
$\tau_{ep}, \tau_{pe}, \tau_{es}$	electron-phonon, phonon-electron, electron-substrate energy transfer time
$\tau_p$	pulse length
$\omega, \omega_1, \omega_2, \omega_3$	angular frequency
$\Psi_n$	$n^{\text{th}}$ Gaussian beam mode
$\Psi_{mn}(x, y, z)$	Gaussian–Hermite beam mode amplitude (rectangular coordinates)
$\Psi_{pm}(r, \phi, z)$	Gaussian–Laguerre beam mode amplitude (cylindrical coordinates)
$\Psi_n^{\text{SC}}$	$n^{\text{th}}$ scattered Gaussian beam mode
$\psi_a, \psi_b, \psi_F$	scalar field
$\psi_G$	scalar Gaussian field

## The THz Band

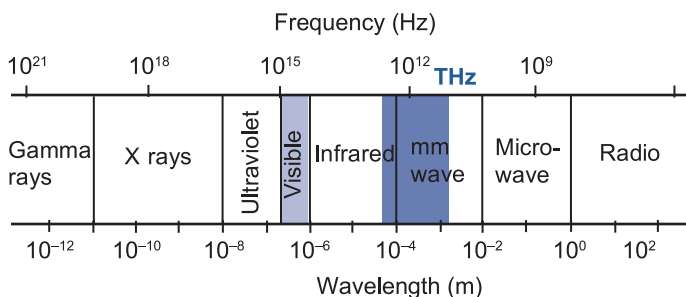
Terahertz (THz,  $10^{12}$  Hz) radiation, alternatively referred to as **T rays**, is electromagnetic radiation lying between the microwave and infrared portions of the spectrum.

There is no standard definition for the **THz band**, but it has most often come to refer to frequencies in the range of 0.3 THz (300 GHz)–3 THz ( $\lambda = 1$  mm–100  $\mu\text{m}$ ), although sometimes it refers to 0.1–10 THz.



## The THz Gap

The lack of readily available THz sources and detectors has led to the relatively unexplored region of the electromagnetic spectrum between microwaves and infrared being termed the “**THz gap**.”



While photon energies at THz frequencies are below the bandgap energies of materials commonly used in solid-state devices or the energy difference in atomic levels used in many lasers, THz frequencies are still difficult to achieve with electronic devices. Nevertheless, in recent years, this gap has been narrowed by microwave electronics on the low-frequency side and photonics on the high-frequency side. In the THz region, where  $h\nu \approx k_B T$  at room temperature, both quantum and thermal noise in devices must be dealt with. (Here  $\nu$  is frequency,  $T$  is temperature,  $h$  is Planck's constant, and  $k_B$  is the Boltzmann constant.)

Although THz radiation is typically propagated and analyzed as free-space beams, unlike in traditional optics, beams may be only a few wavelengths in diameter, and diffraction effects can become important. Neither **geometrical optics**, commonly used in the visible, nor the accurate but computationally intensive **physical optics** techniques used for radio and microwave systems is entirely satisfactory in the THz band.

## THz Absorption in Air

---

**Absorption** and **dispersion** (propagation speeds that depend on frequency) occur as a result of the interaction of electromagnetic radiation with charged particles (nuclei and electrons) in a medium.

In particular, one factor often making research at THz frequencies difficult is the very high absorption of THz radiation by the Earth's atmosphere over much of the band. In this case, the absorption is due largely to rotational transitions of constituent molecules, in particular, water vapor. The relatively low density of molecules gives rise to sharp resonant absorption peaks, separated by transmission windows.

The attenuation from absorption can be approximated using the **Beer-Lambert Law** for **transmission**:

$$\tau = \frac{I}{I_0} = \exp(-\alpha x)$$

where  $I_0$  and  $I$  are the incident and transmitted intensity, respectively,  $x$  is path length, and  $\alpha$  is the absorption coefficient. The absorption coefficient can in turn be written as a product of the absorption cross-section  $\sigma$  and number density  $n$  of absorbers  $\alpha = \sigma n$ .

**Absorbance** is defined as

$$A = -\log_{10} \left( \frac{I}{I_0} \right) = \frac{\alpha x}{\ln(10)}$$

and the depth at which the intensity is reduced to  $1/e$  of the initial value, the penetration depth, is  $1/\alpha$ .

Attenuation can also occur through scattering in the atmosphere, but, because of its longer wavelength, the scattering of THz radiation by airborne particles (dust or fog, for example) is lower than for visible or infrared radiation.

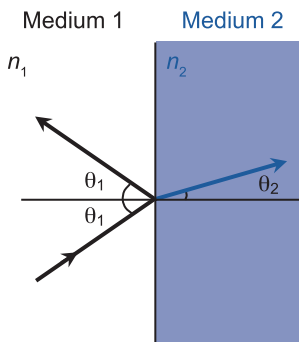
## Reflection and Transmission

When a THz electromagnetic wave is incident upon the boundary between two linear dielectric media (of refractive indices  $n_1$  and  $n_2$ ), the ratios of the reflected and transmitted field amplitudes to the incident amplitude are given by

(s-polarization)

$$\frac{E_{s,r}}{E_{s,i}} = \frac{n_1 \cos \theta_1 - n_2 \cos \theta_2}{n_1 \cos \theta_1 + n_2 \cos \theta_2}$$

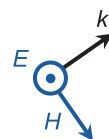
$$\frac{E_{s,t}}{E_{s,i}} = \frac{2n_1 \cos \theta_1}{n_1 \cos \theta_1 + n_2 \cos \theta_2}$$



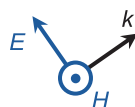
(p-polarization)

$$\frac{E_{p,r}}{E_{p,i}} = \frac{n_2 \cos \theta_1 - n_1 \cos \theta_2}{n_2 \cos \theta_1 + n_1 \cos \theta_2}$$

$$\frac{E_{p,t}}{E_{p,i}} = \frac{2n_1 \cos \theta_1}{n_2 \cos \theta_1 + n_1 \cos \theta_2}$$



s-polarization



p-polarization

where the angles of incidence and refraction are given by Snell's law:

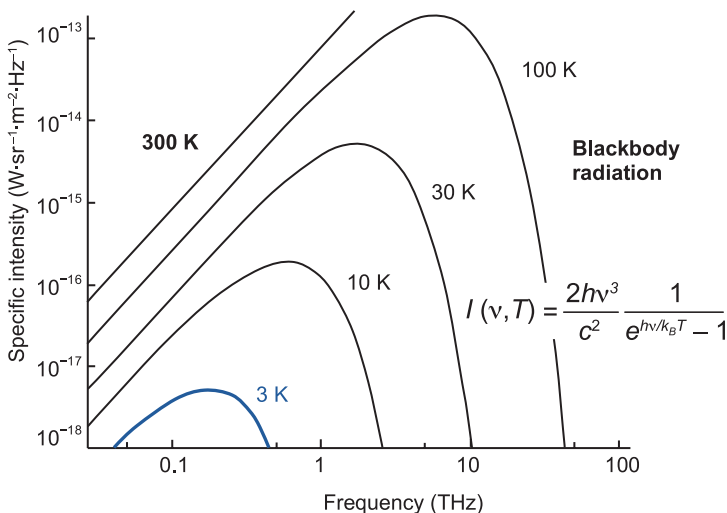
$$n_1 \sin \theta_1 = n_2 \sin \theta_2$$

The **reflectivity**  $R$  and **transmittance**  $T$  are defined by the power ratios

$$R = \frac{|E_r|^2}{|E_i|^2} \text{ and } T = \frac{n_2 \cos(\theta_2) |E_t|^2}{n_1 \cos(\theta_1) |E_i|^2}$$



## Natural Sources of THz Radiation



Electromagnetic radiation, at all frequencies, is emitted by all objects in the universe. This thermal emission is well described by a theoretical **blackbody** spectrum and depends on the temperature of the object. While for natural sources thermal emission is very weak at THz frequencies, cold objects such as 10- to 100-K dust and gas clouds in galaxies emit most of their radiation in this band. The remnant radiation from the Big Bang, known as the cosmic microwave background radiation, has now cooled to a temperature of just below 3 K, and its emission spectrum therefore peaks in the low-frequency end of the THz band.

In addition to broadband thermal emission, the electromagnetic spectrum at THz frequencies is rich in molecular and atomic fine structure lines and contains emission from important species such as CO and OH.

Although natural sources of THz emission are of great importance for passive imaging, in astrophysics in particular, most practical applications involve the use of laboratory-generated THz radiation.

## THz Generation Techniques

---

**Negative differential resistance** devices: If a resonant circuit is combined with a device that exhibits negative differential resistance ( $dV/dI < 0$ ) at approximately its resonance frequency, then oscillations can be sustained indefinitely. The frequency of oscillation is determined by the resonant circuit,  $\nu = 1/2\pi\sqrt{1/LC}$ . In effect, the negative resistance of the device cancels that of the rest of the circuit. The resonant circuit can take the form of a waveguide or cavity resonator. Two-terminal solid state devices such as Gunn diodes, IMPATT diodes, and resonant tunnel diodes are commonly used to generate GHz and THz radiation.

**Lasers:** Lasers are based on the process of stimulated emission in a gain medium. Emission takes place due to a transition from a high-energy to a low-energy state. These states can be, for example, molecular rotational levels (far-infrared gas lasers), conduction subbands (quantum cascade lasers) or Landau levels (p-type germanium lasers) in semiconductors. A pump source maintains population inversion, and the emission is contained within a partially transmissive cavity to allow for amplification. (It is easier to consider the free-electron laser as an electron accelerator source rather than as a conventional laser.)

**Electron acceleration:** Electromagnetic radiation is generated by accelerating electrical charges. Since acceleration generally depends on the charge-to-mass ratio of a particle, THz sources are based on the acceleration of electrons. In some cases these are free electrons in a vacuum (synchrotrons, gyrotrons, free-electron lasers, and backward-wave oscillators); in others they are free charge carriers in photoconductors (photoconductive antennas, photomixers), metallic gratings (Smith–Purcell emitters), or bound charges in semiconductors (difference frequency generators, electro-optic crystals, optical parametric oscillators).

## THz Generation Techniques (cont.)

---

Efficient energy transfer between a free-electron beam and the electromagnetic field occurs if the velocity of the electrons is close to the phase velocity of the electromagnetic field. Fast-wave electron tubes operate at very high voltages to generate the electron velocities required for resonance conditions; conventional slow-wave devices typically contain structures designed to reduce the phase velocity of the field. In order to produce pulsed coherent radiation, electron bunches, comparable in size to the radiation wavelength, are used.

Crystals, in which bound electrons are induced to oscillate by an input electromagnetic field (a nonresonant process), are usually described in terms of the resulting polarization  $P$ , and  $E_{THz} \propto \partial^2 P / \partial t^2$ . In nonlinear crystals the time variation of the polarization, and therefore the frequency of the output field, has components at frequencies other than the input frequency; therefore, they are used for difference frequency generation and optical rectification (sometimes referred to as **frequency down-conversion**).

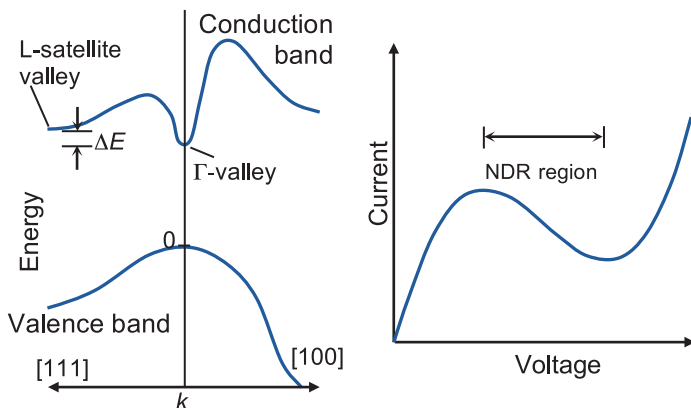
In the case of **photoconductors**, the charge carrier density increases rapidly in response to the intensity envelope of a femtosecond (fs) optical pulse (photoconductive antenna) or the beating of two optical beams closely spaced in frequency (photomixing) and  $E_{THz} \propto \partial J / \partial t$ . The energy of the incoming optical photons must be sufficient to excite carriers across the semiconductor bandgap (called a resonant process because of the photon absorption). Techniques involving femtosecond laser pulses are often known as **ultrafast optics**.

**Frequency multiplication:** In addition to down-converting optical frequencies, THz radiation can be produced by up-converting microwave frequencies in nonlinear devices (frequency multipliers). For maximum efficiency this is generally done in a cascaded series of  $\times 2$  and  $\times 3$  multipliers.

## Gunn Diodes

**Gunn diodes** are two-terminal negative differential resistance (NDR) devices that, when coupled to a suitably tuned ac resonator, generate RF power. Typically, a Gunn diode consists of a uniformly doped n-type III-V material (e.g., GaAs, InP) sandwiched between heavily doped regions at each terminal.

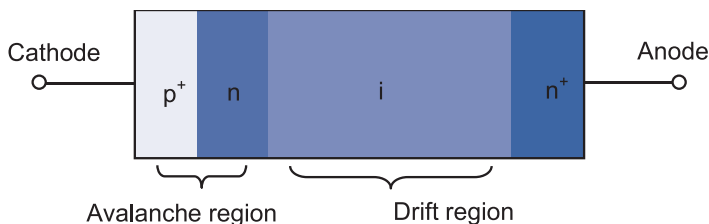
The lowest conduction band in, for example, a GaAs crystal has distinct valleys in certain orientations, two of which are labeled  $\Gamma$  and L. If they reside in the lower  $\Gamma$ -valley, electrons exhibit a small effective mass and very high mobility, whereas in the L-valley, they have a large effective mass and low mobility. The two valleys are separated by a small energy gap  $\Delta E$ .



As the bias voltage across the diode is increased, electrons gain sufficient energy to be transferred to the L-valley. (A Gunn diode is also known as a **transferred-electron device**.) These electrons have a lower drift velocity because of the increase in their effective mass; the current decreases with increasing bias voltage, and the diode exhibits a region of NDR. If biased in this region, small local perturbations in the net charge give rise to ac current oscillations at the contacts. If the diode is placed in a cavity or resonant circuit so that its negative resistance cancels the resistance of the resonator, then the circuit oscillates without attenuation and emits electromagnetic radiation.

## IMPATT Diodes

An **impact ionization avalanche transit time** (or **IMPATT**) **diode** is an NDR device with relatively high power capability. Its basic structure is based on a reverse-biased p-n junction and an intrinsic (high-resistivity) layer.

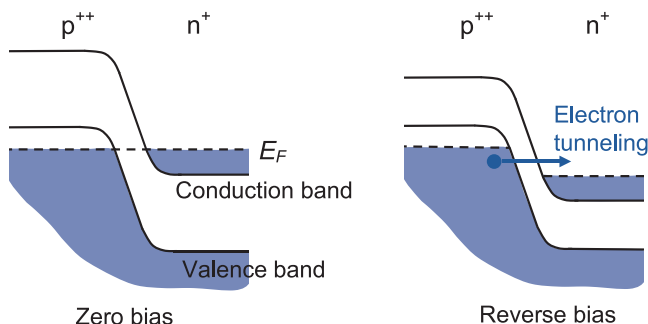


When a reverse bias across the p-n junction exceeds a certain threshold value, then avalanche breakdown (due to impact ionization) occurs, resulting in a large number of carriers in what is called the avalanche region. In the example above, electrons move through the drift region (intrinsic doping) to the n<sup>+</sup> contact in a time known as the transit time delay.

An ac signal with a mean value just below avalanche breakdown is applied to the diode. As the voltage increases above threshold, carriers are generated; however, since the rate of carrier generation at avalanche depends not only on the electric field but also on the number of carriers already in existence, the current continues to rise even as the ac voltage decreases (it lags the voltage by 90 deg, known as the injection delay). The length of the diode can be chosen so that the transit time delay results in a further 90-deg phase lag in current and therefore a negative differential resistance device. An external resonant circuit can then be used to sustain the oscillations. Like other devices based on the random avalanche process, IMPATT diodes tend to suffer from phase noise.

## TUNNETT Diodes

As the operating frequency of an IMPATT diode increases, the dominant carrier injection mode changes from an avalanche to a mixed avalanche–tunnel mechanism. The exploitation of the injection of charge by tunneling led to the development of **tunnel injection transit time** or **TUNNETT diodes**.



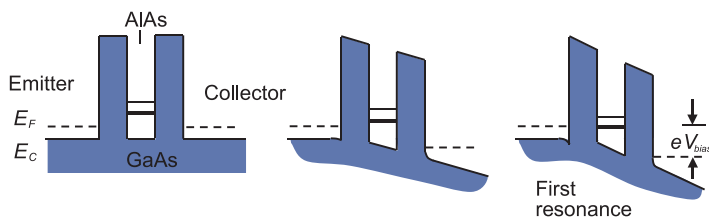
In these diodes a highly doped, narrow,  $p^{++}n^{+}$  junction (in an overall  $p^{++}n^{+}nn^{+}$  structure) is formed in order to change the breakdown mechanism from avalanche to tunnel injection. Since the carrier generation rate due to tunneling does not depend on the current density, there is no injection delay in this case, and negative differential resistance is achieved by the transit time of carriers through a drift region with a considerably lower electric field than in IMPATT diodes.

Tunneling is a quick, low-noise process when compared with impact ionization, so TUNNETT diodes can provide medium power at higher frequencies and with lower noise than IMPATT diodes.

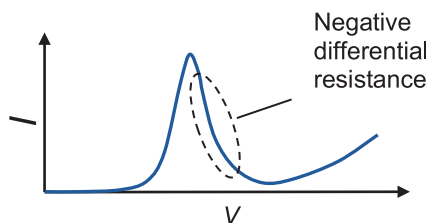
Like other NDR devices, TUNNETT diodes in an external resonant circuit can be used to generate electromagnetic radiation.

## Resonant Tunnel Diodes

In a **resonant tunnel diode (RTD)**, layers of undoped material are used to create a quantum well between two thin barriers. Quasi-bound, or resonant, energy states are formed in the well.



When an energy level in the quantum well is close to the energy of the electrons in the conduction band, then resonant tunneling through the double-barrier structure occurs. This gives rise to a peak in the current-voltage characteristic. The width of the peak depends on the width of the resonant state in the quantum well.



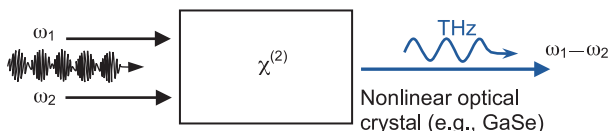
As the bias voltage is increased beyond the resonance peak, the  $I$ - $V$  curve exhibits negative differential resistance.

Like other NDR devices, RTD diodes in an external resonant circuit can be used to generate electromagnetic radiation.

Because tunneling is inherently a very fast process, RTDs produce the highest oscillation frequencies of the NTD devices discussed here.

## Difference Frequency Generation

**Difference Frequency Generation (DFG)** refers to a process in which two beams, at frequencies  $\omega_1$  and  $\omega_2$ , interact in a medium with a second-order nonlinear **susceptibility**  $\chi^{(2)}$ , producing radiation at their difference frequency  $\omega_1 - \omega_2$ .



An electromagnetic field incident on a medium induces bound electrons to oscillate about their equilibrium position. In the linear regime, the resulting dielectric polarization is proportional to the applied electric field

$$\mathbf{P}(t) = \epsilon_0 \chi \mathbf{E}(t)$$

where  $\chi$  is the electric susceptibility, but in the nonlinear regime, when the input field is strong, for example, the polarization is described by a series of terms:

$$\mathbf{P}(t) = \epsilon_0 \chi \mathbf{E}(t) + \epsilon_0 \chi^{(2)} \mathbf{E}^2(t) + \epsilon_0 \chi^{(3)} \mathbf{E}^3(t) + \dots$$

The second-order nonlinear susceptibility tensor  $\chi^{(2)}$  is nonzero in noncentrosymmetric materials; DFG makes use of this second-order term. (For simplicity we consider a lossless dispersionless medium and scalar quantities.)

If the incident optical field has two distinct frequency components  $E(t) = E_1 e^{-i\omega_1 t} + E_2 e^{-i\omega_2 t} + \text{c.c.}$ , then the second-order induced polarization  $P^{(2)}(t) = \epsilon_0 \chi^{(2)} E^2(t)$  has five, at frequencies 0 (**optical rectification**),  $2\omega_1$ ,  $2\omega_2$  (**second-harmonic generation**),  $\omega_1 + \omega_2$  (**sum-frequency generation**), and  $\omega_1 - \omega_2$  (difference-frequency generation).

The input frequencies are chosen so that an output THz radiation field is produced by the time-varying second-order polarization  $E_{\text{THz}}(t) \propto \partial^2 P^{(2)}(t) / \partial t^2$ .



## Electro-optic Crystals (Optical Rectification)

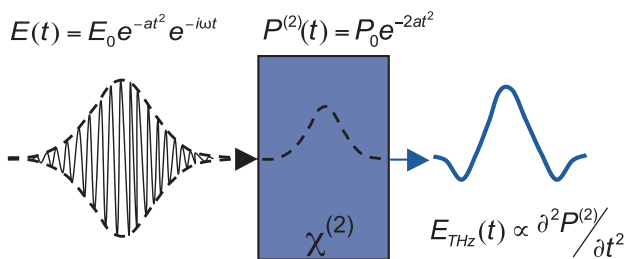
THz radiation can be generated via **optical rectification** of femtosecond optical pulses incident on an **electro-optic crystal**.

If a field with a single frequency component,  $E(t) = E_0 e^{-i\omega t} + \text{c.c.}$ , is incident upon a crystal with a nonzero second-order susceptibility  $\chi^{(2)}$ , then, as before, the nonlinear polarization created is given by

$$P^{(2)}(t) = \varepsilon_0 \chi^{(2)} E^2(t) = 2\varepsilon_0 \chi^{(2)} \left( EE^* + \frac{1}{2} E_0^2 e^{-i2\omega t} \right) + \text{c.c.}$$

The first term describes the process known as optical rectification. In this case it produces a static polarization and so does not lead to the generation of radiation. (The second term refers to second-harmonic generation, which is not used here.)

If, however, the input optical field is a pulse  $E(t) = E_0 e^{-at^2} e^{-i\omega t}$ , then the induced second-order polarization has a time dependence  $P^{(2)}(t) = P_0 e^{-2at^2}$  (i.e., it is proportional to the envelope of the input optical pulse). This time-varying polarization acts as a source of THz radiation.



## Electro-optic Crystals (Optical Rectification) (cont.)

In a nondispersive medium the optical and THz refractive indices are the same, so the incident pulse and the generated THz pulse travel at the same speed. The induced polarization is in phase with the THz radiation amplitude, which grows linearly with distance as a result. In this case extremely high-bandwidth pulses can be generated, limited only by the duration and bandwidth of the optical pulse.

In practice, however, nonlinear crystals are dispersive, and polarization contributes destructively after some distance. For efficient THz generation, the nonlinear material should be shorter than the **walk-off length**

$$l_w = c\tau_p / (n_{THz} - n_{optical})$$

where  $\tau_p$  is the pulse duration, and  $n$  is the refractive index. Velocity matching (or **phase matching**) in a dispersive medium can be achieved if the optical pulse envelope (group velocity) travels at the phase velocity of the THz radiation. ZnTe crystal achieves this best in the case of femtosecond pulses from an 800-nm (**Ti:sapphire**) laser. Alternatively, phase matching has been achieved by novel techniques such as tilting the EO crystal or the use of periodically inverted structures (quasi-phase-matching).

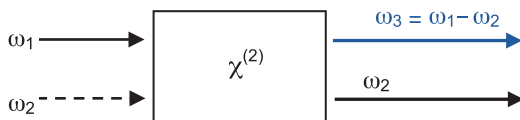
The degree to which the velocity matching condition is met can be expressed in terms of the **coherence length**  $l_c$ , the propagation distance after which the optical pulse leads or trails the THz pulse by a phase of  $\pi/2$ :

$$l_c = \frac{c}{2\nu_{THz} |n_{optical} - n_{THz}|}$$

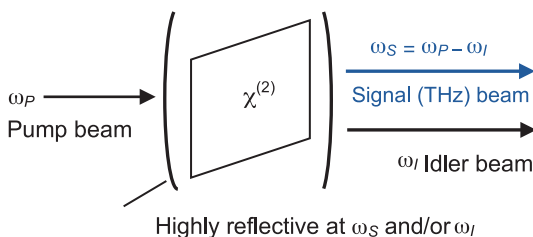
Thin crystals must be used in order to produce broadband THz pulses, at the expense of the THz radiation amplitude.

## Optical Parametric Oscillators

**Optical parametric oscillators (OPOs)** use a nonlinear crystal to convert an input **pump** laser beam into two lower-frequency beams known as the **signal** and **idler** beams.



If DFG is considered in terms of energy conservation, for every photon created at the difference frequency  $\omega_3 = \omega_1 - \omega_2$  (here the THz beam), a photon at the higher input frequency  $\omega_1$  is destroyed, and a photon at the lower input frequency  $\omega_2$  is created. DFG therefore amplifies the lower input frequency beam (idler  $\omega_2$ ), so is known as **parametric amplification** or, if the lower frequency is absent (quantum fluctuations can start the process), **parametric generation**.



The frequencies of the signal and idler beams are determined by the set of frequencies (and bandwidth) for which phase matching ( $\mathbf{k}_P = \mathbf{k}_S + \mathbf{k}_I$ ), and therefore efficient energy transfer, occurs. This in turn can be set by the angle that the crystal makes with the input beam. Other techniques for phase matching include changing the nonlinear crystal temperature or poling period.

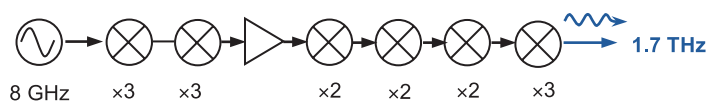
The crystal is typically placed in a cavity that is resonant at either the idler or both signal and idler frequencies.

## Frequency Multipliers

**Frequency multipliers** are nonlinear devices that generate harmonics of an input (microwave in this case) signal and suppress undesired ones. This is analogous to harmonic generation in nonlinear crystals described earlier.

In principle, any electronic device that has a nonlinear response contains higher harmonics of an input signal, although conversion efficiencies vary. Because of the very high switching rates achievable with a metal–semiconductor junction, **Schottky barrier diodes** (described in more detail on page 35) are generally used to upconvert the output frequency of microwave sources, such as Gunn diodes, into the THz range. Either the nonlinear capacitance (**varactor diode**) or the nonlinear resistance (**varistor diode**) of the device is exploited. Multipliers operating in varactor mode have higher theoretical efficiency but lower bandwidth than those operating in varistor mode.

The maximum conversion efficiency of a  $\times N$  multiplier decreases as  $1/N^2$ , while the phase noise increases as  $N$ . For this reason multiplication is typically limited to  $N = 2$  or  $3$ . Multiplier chains with cascaded doublers and triplers can be used to upconvert signals from GHz to THz frequencies.

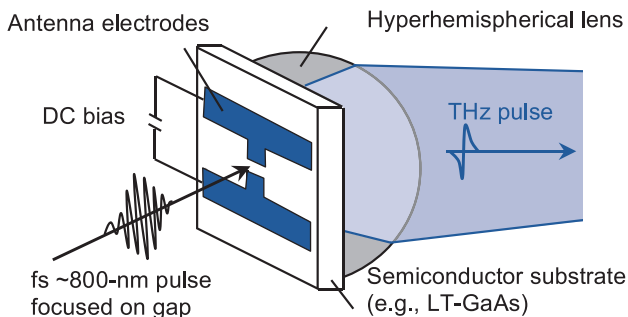


Balanced multiplier configurations that employ a pair of individual multiplier circuits can be used to both add signals at the desired harmonic, increasing output power, and cancel signals at unwanted harmonics, thereby reducing filtering requirements.

## Photoconductive Antennas

Broadband THz pulses can be produced from the transient photocurrent in semiconductors induced by ultrashort optical pulses from femtosecond lasers.

The ultrashort optical pulses are produced by mode-locked femtosecond lasers, often with a gain medium of titanium-doped aluminum oxide (Ti:sapphire). A typical Ti:sapphire laser produces 800-nm pulses of 10- to 100-fs duration at a repetition rate of  $\sim 80$  MHz.



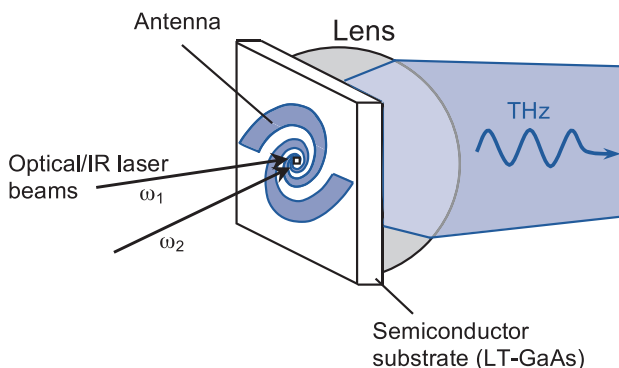
A **photoconductive antenna (PCA)** operates as an electrical switch by exploiting the increased electrical conductivity of a semiconductor when exposed to light. As long as they have an energy greater than that of the semiconductor's bandgap energy, the incoming photons generate free carriers in the semiconductor and allow a photocurrent to flow.

The PCA consists of two dc-biased electrodes on a semiconductor substrate. Optical pulses generate free carriers, which are then accelerated by the bias across the gap before becoming trapped in semiconductor defect sites. If this occurs on a subpicosecond timescale (this depends on the optical pulse duration and the carrier lifetimes), the resulting impulse current generates a THz pulse.

Commonly used electrode antenna structures include the stripline, dipole, and bowtie. A substrate lens can be attached to the back of the PCA to refocus the divergent THz beam.

## Photomixing

**Photomixing** (or **optical heterodyning**) involves the use of two infrared or visible lasers to generate THz radiation at their difference (beat) frequency. The mixing of the two input signals is generally done in a biased antenna-coupled **photoconductive switch**.



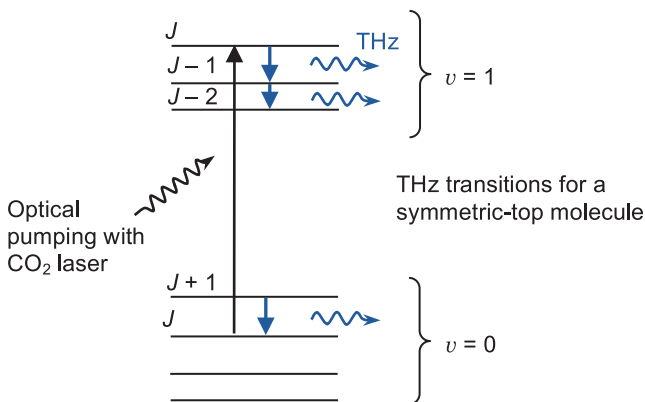
Two CW lasers, typically operating at 800–850 nm, are overlapped in space and focused onto a biased region of a semiconductor photoconductor such as low-temperature-grown gallium arsenide (LT-GaAs). The intensity of the light is modulated at the beat frequency of the two sources and, providing its energy is greater than the semiconductor bandgap energy and the carrier lifetime of the semiconductor is sufficiently short, it induces a photocurrent at the same frequency,  $\omega_1 - \omega_2$ . The current then generates THz radiation.

A series of closely spaced electrodes are usually connected to a patterned antenna structure to improve the coupling to free space. The radiation can be further focused by placing a hyperhemispherical silicon lens on the semiconductor substrate.

The output powers of photomixers tend to be low ( $\mu\text{W}$ ) because the low conductivity of LT-GaAs limits the level of possible continuous optical excitation.

## Optically Pumped Far-IR Gas Lasers

In **optically pumped far-IR gas lasers** [sometimes called **optically pumped THz lasers (OPTLs)**], THz emission originates from transitions between rotational states of gas molecules that have permanent electric dipole moments, such as  $\text{CH}_3\text{F}$ ,  $\text{CH}_2\text{F}_2$ , or  $\text{CH}_3\text{OH}$ .



Optical pumping with, in most cases, a  $\text{CO}_2$  laser excites some molecules from the lowest ( $v = 0$ ) to the first excited vibrational state ( $v = 1$ ). The resulting population inversion between rotational states, corresponding to the excited vibrational level, gives rise to THz radiation. The exact transition selection rules involved depend on the type of gas molecule.

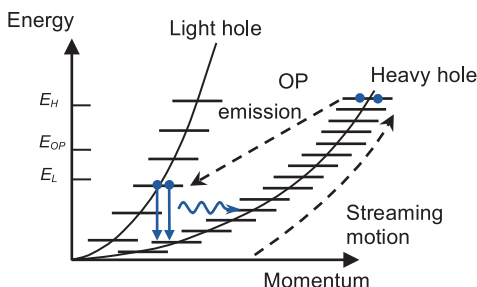
Because of the large energy difference between the rotational and vibrational energy level separations, the complete process is quite inefficient ( $\nu_{\text{THz}}/2\nu_{\text{pump}}$ ) with the majority of the pump radiation converted to heat. Nevertheless, OPTLs can produce powers of about 100 mW at frequencies up to 8 THz.

The output THz radiation can be polarized in a direction orthogonal or parallel to the pump radiation, depending on the transition.

## p-Type Germanium Lasers

The **p-type germanium (p-Ge) laser** is a tunable, solid state laser based on population inversion between light- and heavy-hole valence subbands in crossed electric and magnetic fields.

In Ge, the top of the valence band consists of degenerate bands. In p-type Ge, holes have different effective masses in two of these bands and are known as the light- and heavy-hole valence bands. In the presence of a magnetic field, the bands have their energies quantized into discrete levels known as **Landau levels** (quantization of cyclotron orbits).



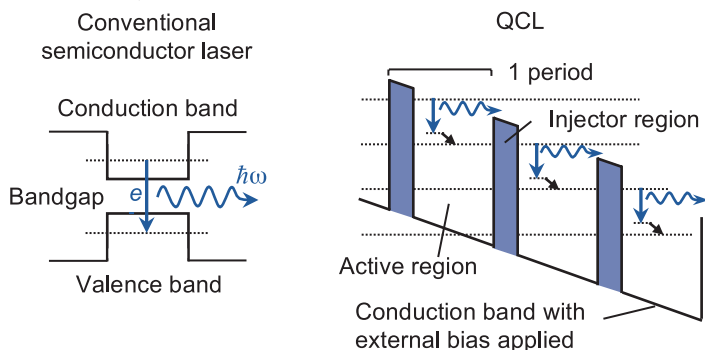
If an electric field  $\mathbf{E}$  is applied perpendicular to the magnetic field  $\mathbf{B}$ , then heavy holes can be accelerated in a streaming motion to reach the optical phonon (OP) energy  $E_{OP}$  and above. Light holes reach a different maximum energy  $E_L$  and the ratio of  $E/B$  can be adjusted so that  $E_L < E_{OP}$ . In this case, heavy holes can spontaneously emit phonons and scatter into the light-hole band, but the light holes cannot emit phonons. This leads to an accumulation of holes in relatively long-lived light-hole Landau levels (population inversion). Laser action occurs due to direct optical transitions between the light- and heavy-hole subbands over a wide frequency range.

p-Ge lasers are tunable over about 1–4 THz by changing the applied electric/magnetic fields. They are operated at liquid-helium temperatures in order to avoid optical and acoustic lattice scattering, but heat loading in the Ge crystal still limits the laser to relatively short pulse lengths and slow repetition rates.



## Quantum Cascade Lasers

**Quantum cascade lasers (QCLs)** are semiconductor heterostructures that can emit CW THz radiation. Unlike conventional interband semiconductor lasers that emit photons through the recombination of electrons from the conduction band with holes from the valence band, QCLs are unipolar (electron-only) devices that exploit transitions between conduction band states (intersubband transitions).

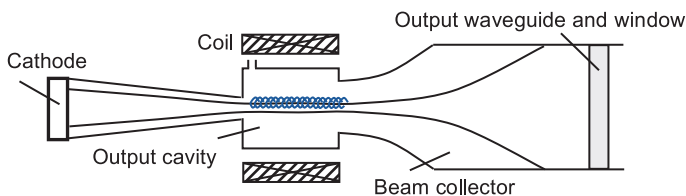


The conduction band offsets between neighboring materials in alternating layers (the superlattice) create a series of quantum wells and barriers. Quantum confinement leads to the splitting of the conduction band energy levels into subbands; QCLs are generally based on a three-layer system. Electrons are injected into level 3 of the active region, and then the laser transition occurs to level 2, followed by rapid depletion to level 1. The injector region collects the electrons and injects them into level 3 of the next period. Active regions are usually made of several quantum wells, and a suitable design of layer thicknesses creates a population inversion between levels 2 and 3. The cascading of superlattice periods means that one electron can produce many photons. The most widely used QCL designs are based on GaAs/AlGaAs superlattices and are known as **chirped superlattice**, **bound-to-continuum**, and **resonant phonon** designs.

Current challenges in developing THz QCLs are to reduce the minimum operation frequency, increase the frequency tuning range, and increase the operation temperature.

## Gyrotrons

**Gyrotrons** (sometimes called **cyclotron resonance masers**) are fast-wave electron tube devices that are used when high-power coherent radiation is required. A gyrotron typically contains an annular electron beam that is accelerated to (weakly) relativistic speeds before being introduced into a resonance cavity that is immersed in a strong magnetic field. The electrons gyrate in the longitudinal magnetic field at their cyclotron resonance frequency given by  $\omega = eB/\gamma m_e$ , radiate, and then interact with the electromagnetic field in the cavity. The only requirement on the cavity is that it supports a TE mode, at a frequency slightly higher than  $\omega$ , with a strong azimuthal electric field at the electron beam radius.



The relativistic electrons are initially randomly distributed in terms of the angular position (phase) in their cyclotron orbit, and there is no net energy transfer to the electromagnetic field (some are accelerated by the electric field, some decelerated). However, **phase bunching** occurs for a wave frequency that is slightly greater than the initial value of the cyclotron frequency. (This phase bunching occurs because the electron's relativistic mass, and therefore its gyration frequency, changes as it gains or loses energy.) Once bunched, the electrons radiate coherently, and a large fraction of the electron beam energy is converted to electromagnetic radiation, which is then coupled out of the cavity.

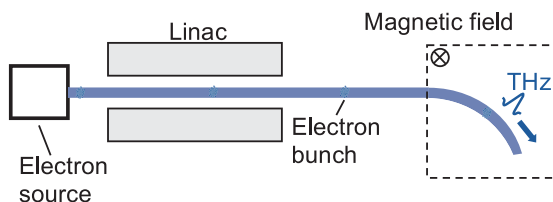
The cyclotron frequency, and therefore the frequency of the radiation emitted, is determined by the strength of the magnetic field and not by the dimensions of a resonant cavity. For THz radiation, strong magnetic fields are needed, or alternatively, it is possible to obtain some power at harmonics of the fundamental frequency.

## Synchrotrons

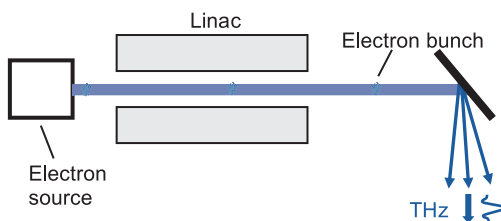
An accelerating charged particle emits broadband electromagnetic radiation. In the case of a relativistic electron undergoing circular motion (in a magnetic field, for example), the radiation is called **synchrotron radiation**, and it is highly beamed into a narrow cone ( $\Delta\theta = 1/\gamma$ ) in a direction perpendicular to the motion of the electron.

In order to produce pulsed coherent radiation, electron bundles—comparable in size to the radiation wavelength—are used. In this case, the intensity of the radiation produced is proportional to the square of the number of electrons and can be very high.

Linear accelerators (**linacs**) and **storage rings** can be used to produce synchrotron radiation. The free-electron laser, discussed next, is another example of a synchrotron source.



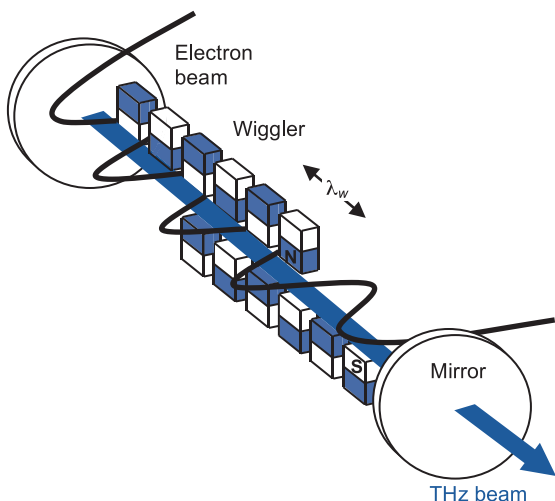
Electron accelerators can also be used to produce **bremmstrahlung** radiation by accelerating the electrons toward a fixed metal target.



At present, synchrotron sources are provided as part of large national or international facilities.

## Free-Electron Lasers

In a **free-electron laser (FEL)**, a series of magnets, called the **wiggler** or undulator, produces a periodic magnetic field that is transverse to the direction of a relativistic beam of electrons.



The electrons are forced to undergo sinusoidal oscillations, emitting **synchrotron radiation** in the forward direction as they accelerate. If the wiggler is placed inside a low-loss cavity, then the interaction between the electrons and the radiation field (velocity modulation) causes the electrons to bunch at intervals of the radiation wavelength, and the radiation emitted is coherent.

FELs operate in pulsed mode. The wavelength of the emitted radiation is given by

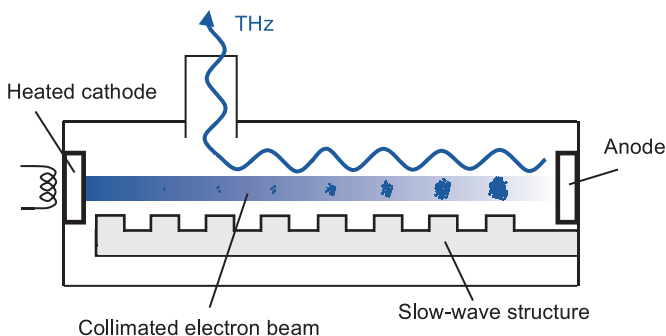
$$\lambda_{FEL} = \frac{\lambda_w}{2\gamma^2} (1 + K^2), \text{ with } K = \frac{eB\lambda_w}{2\pi m_e c^2}$$

An advantage of the FEL over a conventional laser is that it can be continuously and widely tuned by adjusting the energy of the electron beam ( $\gamma m_e c^2$ ) or the average magnetic field strength  $B$  of the wiggler. FELs produce very high powers but require electron accelerators and so are not practical for standard laboratory use.

## Backward Wave Oscillators

A **backward wave oscillator (BWO)** works by extracting energy from a bunched electron beam to a high-frequency RF electric field. For an effective transfer of energy, the drift velocity of the electrons must be slightly greater than the phase velocity of the electromagnetic field.

A heated cathode at one end of a vacuum tube emits electrons that are collimated by a magnetic field and accelerated toward an anode at the other end. Along the length of the tube there is a periodic slow-wave structure that supports an electromagnetic wave (both forward and backward space harmonics). This electric field modulates the velocity of the electrons, and the electrons become bunched.



The periodic structure is designed so that one component of the electric field, a backward-traveling wave, has the required phase velocity for energy transfer from the electrons. The amplified THz electromagnetic wave is then coupled out through a waveguide.

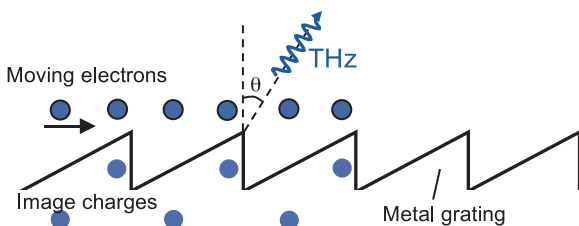
The frequency of generated radiation depends on the velocity of the electrons and therefore the dc bias across the tube. In practice, BWOs can be tuned across  $\pm 10\%$  in frequency. They are generally used below  $\sim 1$  THz.

## Smith–Purcell Emitters

**Smith–Purcell radiation (SPR)** is the broadband radiation emitted when a charged particle, say, an electron, passes above a periodic grating structure. The radiation is angularly dispersed in wavelength according to the dispersion relation

$$\lambda = \frac{l}{|n|} \left( \frac{1}{\beta} - \sin \theta \right)$$

where  $l$  is the grating period,  $\beta c$  is the velocity of the electron where  $c$  is the speed of light,  $n$  is the spectral order, and  $\theta$  is the angle as measured from the surface normal. SPR is generally explained as being due to the acceleration of image charges induced on the surface of the grating by the moving electron.



An induced evanescent wave travels along the grating surface and radiates when it reaches the end of the grating, but at a frequency less than that of the SPR. Coherent SPR can be generated from a bunched electron beam if the bunches are short compared with the radiation wavelength. The electron beam could be either initially bunched or become bunched by the strong interaction of a continuous beam with the surface evanescent wave.

In addition, the bunching of an electron beam at the evanescent wave frequency causes **superradiance** (superlinear dependence of the radiated power on beam current, indicating a stimulated process), which changes the intensity of the SPR, particularly at the bunching frequency and its harmonics. It is the transfer of energy from the electrons to a “slow” evanescent wave—if it has a phase velocity close to the electron velocity—that indirectly drives the radiating modes.

## THz Detection Techniques

---

THz detection schemes are often classified as being either **coherent** or **incoherent**, depending on whether they measure the amplitude and phase of the input field (coherent) or only its intensity (incoherent). Incoherent (or **direct**) detectors are typically broadband.

**Thermal** devices detect CW THz radiation through the increase in temperature that it causes in a sensing element (absorber). The detector exploits some property that changes with temperature, such as electrical resistivity in the case of bolometers, electrical polarization in the case of a pyroelectric detector, and gas pressure in the case of a Golay cell. Thermal detectors generally have a broad spectral response but can be relatively slow, since the absorbing element must reach thermal equilibrium before the measurement is made.

Microwave techniques, such as **rectification**, have been extended to higher frequencies for use in THz detection. Rectifiers, such as Schottky diodes, exploit a nonlinear current–voltage relationship to produce an output voltage that is proportional to the power (amplitude squared) of an input current that has been induced at the THz radiation frequency.

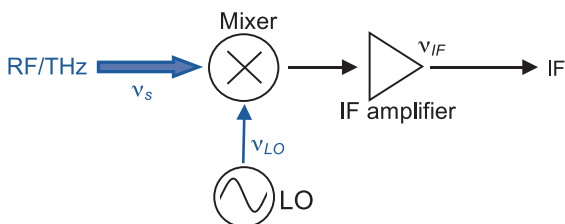
In a process related to optical rectification (the **Pockels effect**), electro-optic sampling senses the electric field of THz radiation from the birefringence it induces in nonlinear optical media.

In **photodetectors** (usually photoconductors), individual signal photons interact directly with electrons in the detector. At THz frequencies, photon energy is small and corresponds to the energy gap between shallow impurity states and the conduction or valence band in extrinsic semiconductors. Incoming THz photons excite electrons from donor states, or into acceptor states, and the change in the number of free carriers produces a measurable change in device resistance.

## THz Detection Techniques (cont.)

Intrinsic photo-excitation of free carriers in a semiconductor generally requires higher-energy (optical) photons. Photoconductive antennas can however be used as THz detectors by allowing the input radiation to induce a current across the photoconductive gap once the photocarriers themselves have been generated by an optical probe pulse. In the case of photomixing the number of photocarriers is modulated by the beat of two optical beams, and the photocurrent depends on the relative phase between the optical beat and the THz radiation (i.e., coherent detection).

**Heterodyning** is a coherent technique that downconverts a THz signal to one at a lower frequency where it can be amplified by standard techniques.



Heterodyning is achieved by mixing the signal to be detected ( $\nu_s$ ) with a stable local oscillator (LO) signal at a nearby frequency ( $\nu_{LO}$ ). The mixer is a device with a nonlinear  $I$ - $V$  characteristic, so that its output contains a component at an intermediate frequency (IF)  $\nu_{IF} = |\nu_s - \nu_{LO}|$ . So long as the LO is stable, this component retains the spectral and phase information of the original signal. Schottky diodes and semiconductor-insulator-semiconductor (SIS) devices are commonly used at frequencies below 1 THz. Hot-electron bolometers, although they are thermal detectors, are sufficiently fast to respond to the IF and can be used as mixers at frequencies above 1 THz.



## Responsivity and Signal-to-Noise Ratio

The performance of detectors can be characterized by a number of different figures of merit. Manufacturers often describe the relationship between the radiation power incident on a detector and its output by the detector's **responsivity**  $R_V$ . Since most detectors produce an electrical output, the responsivity is typically specified as either the voltage or current output per watt of incident power (units V/W or A/W).

In many cases, the incident radiation is modulated and then amplified with a certain amplification bandwidth  $\Delta\nu$ . The responsivity is usually a function of this bandwidth, bias voltage, modulation frequency, etc., so these must be specified when quoting responsivities.

From measurements of both the responsivity and root-mean-square (rms) noise level  $n$ , the **noise equivalent power (NEP)** of a detector can be calculated as

$$NEP = \frac{n}{R_V \sqrt{\Delta\nu}}$$

Another figure of merit is the ratio of detector signal to background noise, or **signal-to-noise ratio (SNR or S/N)**:

$$SNR = \frac{P_{signal}}{P_{noise}}$$

where  $P$  refers to power, and the signal and noise are measured over the same bandwidth. For signals with a large dynamic range, this ratio is often measured in decibels:

$$SNR_{dB} = 10 \log_{10} \frac{P_{signal}}{P_{noise}}$$

In a photon-counting detector,

$$SNR = \frac{n_{signal}}{n_{noise}}$$

where  $n$  refers to the number of photons detected.

## Noise Equivalent Power

**Noise equivalent power (NEP)** is a basic measure of the sensitivity of a detector. For radiation detectors, it is generally defined as the radiant power normalized to a 1-Hz bandwidth, which produces an SNR of unity. It is specified for a given operating wavelength and modulation frequency. When defined in this way, NEP has the units of  $\text{W/Hz}^{1/2}$ .

The advantage of expressing a detector's sensitivity in NEP as opposed to an SNR is that the NEP remains constant for a given detector. The two are related by

$$SNR = \frac{P_{\text{signal}}}{NEP \sqrt{\Delta\nu}}$$

where  $P_{\text{signal}}$  is the signal power incident on the detector, and  $\Delta\nu$  is the integration bandwidth.

Obviously, it is desirable to have as low an NEP as possible, since this indicates the minimum detectable signal ( $SNR = 1$ ) when averaging over a time of the order of one second. Weaker signals can be detected using phase-sensitive detection provided the signal bandwidth is much smaller than the detection bandwidth.

The NEP can be related to the responsivity  $R_V$  of a detector at a specified frequency by

$$NEP = \frac{\text{Noise Voltage}/\sqrt{\text{Hz}}}{R_V}$$

where noise can refer to detector or background noise at the same frequency.

An alternative performance figure of merit is the **detectivity**  $D$  defined as the reciprocal of the NEP. In the case of a Johnson-noise-limited detector, specific detectivity can be used since the noise scales with the square-root of the detector area  $A$ :

$$D^* = \frac{A^{1/2}}{NEP}$$

## Shot and Thermal Noise

**Shot** (or **quantum**) **noise** describes the uncertainty in the number of photons detected in any time period due to the statistical fluctuations in the occurrence of independent random events. It is a quantum noise effect related to the discrete nature of photons and is an inherent property of the incoming light field. (Randomness in the conversion of photons to an electrical signal, for example, can be an additional source of noise.)

The number of detected photons can be described by a Poisson distribution, which has a standard deviation equal to the square root of the average number of photon detections  $n$ ; therefore, the SNR is given by

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

For large numbers, the Poisson distribution approaches a normal (or Gaussian) distribution, and in practice, shot noise is often indistinguishable from Gaussian noise.

In the classical regime, where  $h\nu \ll k_B T$ , or for detectors with a low responsivity, the dominant source of noise is often internal electrical noise (random fluctuations in signal) associated with the detector itself. This **thermal** (also called **Johnson**) **noise** is generated by the random thermal motion of charge carriers and is approximately constant with frequency. For example, the rms voltage due to thermal noise generated in a resistor over a bandwidth  $\Delta\nu$  is given by

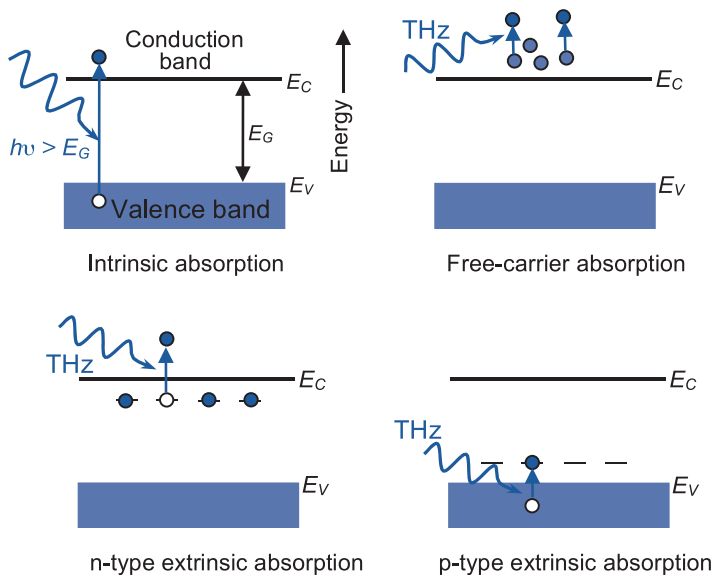
$$V_{rms} = \sqrt{4k_B T R \Delta\nu}$$

where  $T$  and  $R$  are the resistor's temperature and resistance, respectively. Since the amount of generated thermal noise depends on the temperature of the device, thermal noise-limited detectors are often cooled.

## Extrinsic Semiconductor Detectors

Semiconductors have a conduction and a valence band separated by an energy gap  $E_g = E_C - E_V$ . If an incoming photon has an energy greater than the energy gap, then it can excite an electron from the valence to the conduction band (**intrinsic excitation**), thus generating an electron–hole pair. The resulting increase in conductivity can be measured.

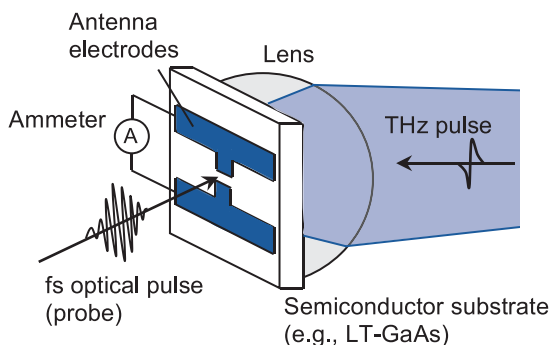
If the photon does not have enough energy to overcome the gap (as is the case for radiation with  $\nu < 10$  THz), then impurities can be added to create donor or acceptor states close to the band edges. Low-energy photons are then sufficient to excite electrons from donor states into the conduction band or from the valence band into acceptor states (**extrinsic excitation**). Detectors are cooled to 4 K or less, and low-impurity levels are used in order to reduce thermal excitation (**dark current**).



Mechanically stressed gallium-doped germanium (Ge:Ga) has been the most widely used semiconductor for sensitive **extrinsic** THz photoconductors.

## Photoconductive Detectors

**Photoconductive antennas (PCAs)** excited by ultra-short (femtosecond) laser pulses can be used to detect as well as generate broadband THz radiation.



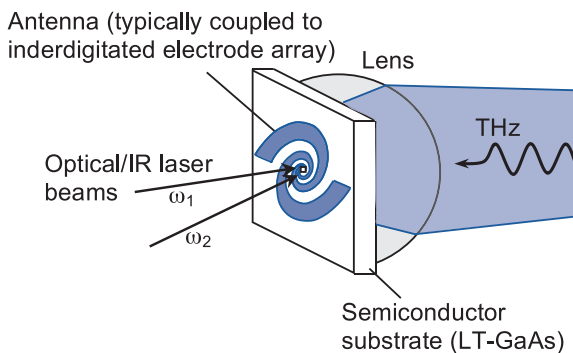
In the case of a **photoconductive detector** (switch), the **THz pulses** are directed onto an unbiased PCA, thereby inducing a time-varying bias across the antenna electrodes. An optical laser pulse (called the **probe**) incident on the gap between the PCA electrodes excites free carriers in the semiconductor. The excited carriers are accelerated by the electric field of the incident THz pulse, leading to a photocurrent that can be amplified and measured in an external circuit. The current is proportional to the amplitude of the incoming THz radiation, convolved with the conductivity that is induced by the femtosecond optical pulse.

Low-temperature-grown GaAs (LT-GaAs) is widely used as the detector substrate because of its very short carrier lifetime, large resistivity, and good carrier mobility. The duration of the photocurrent depends on the carrier lifetime (ps), and since this is generally shorter than the THz pulse, PCAs can be used for time-resolved waveform measurements.

The ultrashort optical pulses are typically produced by mode-locked femtosecond Ti:sapphire lasers (~800 nm).

## Photomixers

Photoconductive antennas (PCAs) that are illuminated by two single-frequency ( $\omega_1, \omega_2$ ) optical lasers can be used for coherent THz detection in both pulsed and CW modes, provided that the difference between the frequencies of the lasers is equal to that of the THz wave.



An incident THz field induces a time-varying voltage across the active region of a photoconductor via some antenna structure.

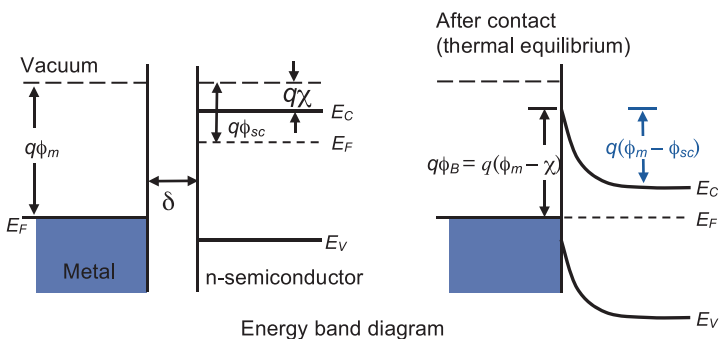
At the same time, the conductivity of the photoconductor substrate (typically LT-GaAs) is modulated at the difference or beat frequency of the two laser beams  $\omega_1 - \omega_2$ . This occurs as long as the energy of the photons is greater than that of the semiconductor bandgap and free carriers are generated. The photocurrent resulting from the acceleration of these carriers by the THz-induced voltage (also modulated at  $\omega_1 - \omega_2$ ) can be measured, and the amplitude of the dc component is proportional to the amplitude of the incident THz radiation. The relative phase between the laser beats and the THz radiation can often be adjusted to maximize this dc component.

**Photomixers** can be used to recover both amplitude and phase information from the incident THz wave.

## Schottky Diodes

**Rectifying diodes** have a nonlinear response to an (usually antenna-coupled) electric field and, for small signals, produce an output that is proportional to the square of the input voltage, i.e., power. They are sometimes referred to as **crystal detectors**.

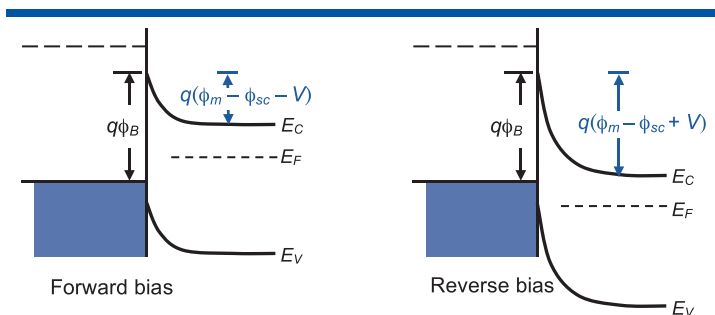
In a **Schottky diode** a metal–semiconductor junction is formed, creating a Schottky barrier at the interface, instead of the usual semiconductor–semiconductor junction in conventional p–n diodes.



In the case of an n-type semiconductor, electrons flow from the semiconductor to the metal, and this equalizes the two Fermi levels. As the negative charge builds up on the metal surface, a corresponding positive charge builds up in a barrier (depletion) region in the semiconductor due to the ionized donor locations left behind. An electric field is set up between these positive charges and the electrons. This field eventually inhibits further electron flow into the metal.

The potential barrier height to electron flow from the metal is equal to the difference between the metal work function ( $q\phi_m$ , where  $q$  here is the electron charge), the electron affinity of the semiconductor ( $q\chi$ ), and the potential developed across the junction, known as the built-in potential  $V_{bi} = \phi_m - \phi_{sc}$ . Electron conduction occurs primarily by **thermionic emission** over the barrier.

## Schottky Diodes (cont.)



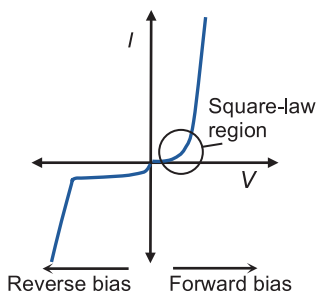
If a positive bias is applied to the metal (forward bias  $V$ ), its Fermi level is lowered with respect to that of the semiconductor. Electron energy is increased relative to the barrier height, allowing increased electron emission from the semiconductor into the metal. A positive current flows through the junction at a voltage comparable to  $V_{bi}$ . A Schottky diode has the typical forward semiconductor diode characteristic, but with a much lower turn on voltage.

If a negative voltage is applied, the potential across the semiconductor increases and almost no current flows resulting in a pronounced rectifying behavior. (The barrier height  $q\phi_B$  as seen from the metal is virtually independent of the bias.)

The  $I$ - $V$  characteristic of the forward-biased diode is given by the ideal diode equation

$$I(V) = I_0 [\exp(qV/k_B T) - 1]$$

where  $I_0$  is the current parameter.



Schottky diodes are used at these frequencies because they provide higher switching speeds than conventional diodes. They are used at room temperature and have responsivities of several hundred V/W.



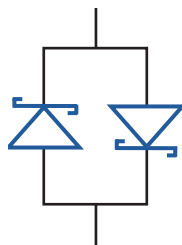
## Schottky Diode Mixers

The nonlinear current–voltage characteristic of the Schottky diode means that it can also be used as the **mixing** element in a heterodyne receiver.

In this type of receiver, the signal that is to be detected is combined in the mixer with a local oscillator signal, and the output contains components at an intermediate difference frequency that is easier to filter and amplify than the original signal.

Along with the low forward voltage drop, the fact that in Schottky diodes only the majority carriers play a significant role in the normal operation of the device allows for much faster transitions between the conducting and nonconducting state when compared with the usual p–n rectifier diode. This makes Schottky diodes suitable for operation as mixers at high frequencies. They are commonly used as mixers at frequencies up to 1 THz, as the large local oscillator powers required are available in this frequency range. The barrier height, a design variable for a Schottky diode, determines the amount of LO power required.

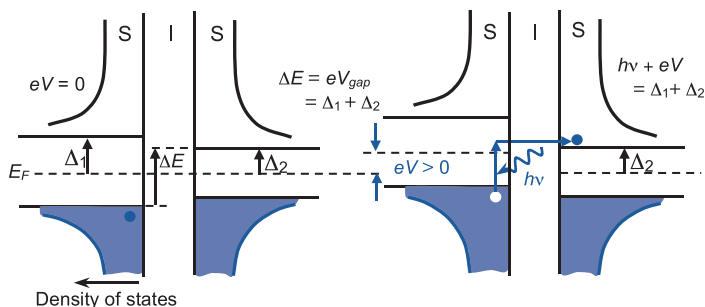
Most mixers use multiple diode techniques to solve problems, such as the coupling of input power to the LO circuit. Different configurations are possible; the subharmonic mixer, for example, uses an antiparallel pair of Schottky diodes to generate harmonics of the LO frequency, thereby allowing a less-expensive lower-frequency LO to be used.



For example, mixing efficiency can be measured by the conversion loss, that is, the ratio of input RF/THz power to output IF power, or by the noise power, which is the ratio of the SNR at the input to the SNR at the output.

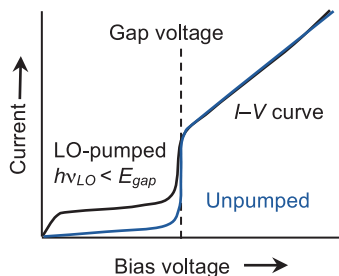
## SIS Mixers

Superconducting tunnel junctions, or **superconductor-insulator-superconductor (SIS)** devices, can be used as the mixer in a heterodyne receiver. **SIS mixers** are extremely sensitive—almost quantum limited.



An SIS mixer consists of two superconducting electrodes separated by a thin insulating layer. If the density of states of the two superconductors have an energy gap of  $2\Delta_1$  and  $2\Delta_2$ , respectively, then at  $T = 0$  and a bias voltage below the gap voltage,  $V_{gap} = (\Delta_1 + \Delta_2)/e$ , no current flows. When the bias voltage is increased to the gap voltage, the Cooper pairs on one side of the junction can break up into two electrons (the quasiparticles), which then tunnel to the other side of the junction before recombining, resulting in a sharp increase in current. (Tunneling of Cooper pairs at zero applied voltage can also occur.)

When an SIS junction is biased below its gap voltage and operated as a mixer in the presence of radiation, a quasiparticle can overcome the energy gap and tunnel through the barrier, partly from the applied bias voltage and partly by absorbing a (THz) photon.



It is the strong nonlinearity in the  $I$ - $V$  characteristics of this **photon-assisted tunneling** of quasiparticles (even at  $T \approx 4$  K) that makes the SIS an excellent mixer.

## Electro-optic Sampling

---

THz radiation can be detected with nonlinear optical crystals by exploiting the Pockels (or **linear electro-optic**) effect.

As before, the dielectric polarization induced in a crystal by a strong input field  $\mathbf{E}(t)$  is described by a series of terms:

$$\mathbf{P}(t) = \varepsilon_0 \chi \mathbf{E}(t) + \varepsilon_0 \chi^{(2)} \mathbf{E}^2(t) + \dots$$

the second of which is nonzero in noncentrosymmetric materials.  $\chi^{(2)}$  is called the second-order nonlinear susceptibility tensor.

The components of the second-order polarization can be written as

$$P_i(\omega) = 2\varepsilon_0 \sum_{jk} \chi_{ijk}^{(2)} E_j(\omega) E_k(0)$$

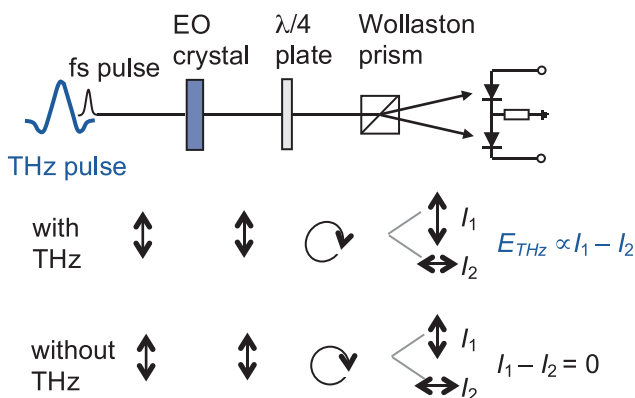
where, in this case,  $E_j$  represents an optical femtosecond pulse, and  $E_k$  the (quasi-static) field of a THz pulse. Here,  $i$ ,  $j$ , and  $k$  represent the Cartesian spatial axes  $x$ ,  $y$ , and  $z$ , respectively.

This change in polarization, or equivalently refractive index, induced by the presence of a static (or low-frequency) electric field, is the Pockels effect. (It is related to the optical rectification discussed on page 13, where  $E_j$  and  $E_k$  would represent the same field resolved along different axes.) If the THz phase velocity matches the optical group velocity, then it can be considered to present a constant electric field to the optical pulse. This occurs near 800 nm (Ti:sapphire laser) for ZnTe crystals. The induced polarization is perpendicular to the optical and THz fields, leading to a field-induced birefringence.

In the presence of a THz field, therefore, a linearly polarized optical probe pulse becomes slightly elliptically polarized as it travels through the crystal.

## Electro-optic Sampling (cont.)

A typical detection setup is shown below. After traveling through the electro-optic crystal, the probe beam passes through a  $\lambda/4$  plate followed by a Wollaston prism, which splits the probe beam into two orthogonal linearly polarized beams. An output signal is obtained by measuring the difference in the intensity of the two beams using a balanced photodetector.



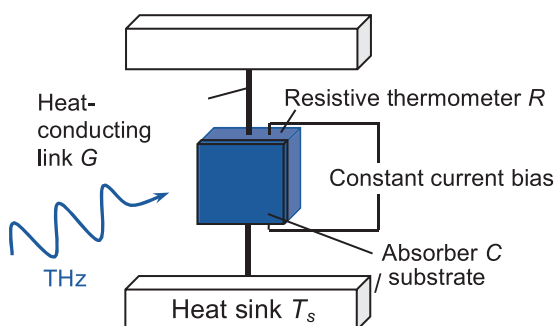
Detection sensitivity is maximized if the Wollaston prism is oriented such that the polarization of the two propagation modes of the plate and the polarization directions of the two beams produced by the Wollaston prism are at 45 deg relative to each other.

The duration of the probe pulse is usually much shorter than the THz pulse, so only a small part of the THz pulse is measured. By varying the relative time delay between the probe and THz pulse, both the amplitude and phase of the THz pulse can be determined.

The spectral resolution of an electro-optic detector is limited by how well the optical and THz velocities match, the frequency dependence of the susceptibility, and by the finite duration of the optical pulse probe.

## Semiconductor Bolometers

**Bolometers** are thermal detectors that exploit the change in electrical resistance of a material with temperature. The material can be a metal, a semiconductor, or a superconductor (TES). Heavily doped semiconductors such as ion-implanted Si or neutron-transmutation-doped (NTD) Ge are commonly used. Bolometers that separate the two main functions, radiation absorption and temperature sensing, are known as **composite bolometers**.



Incoming radiation increases the temperature of the absorber, which is thermally attached both to the thermometer resistor and, via a weak thermal connection, to a heat sink at a low temperature. At low operating temperatures, the correct semiconductor doping level ensures that the dominant conduction mechanism is **electron (variable range) hopping** between donors, thus keeping the free carriers closely coupled to the lattice and ensuring that the temperature of the whole sample, and not just the free carriers, rises. The temperature increase induces a change in the resistance of the semiconductor. This change is sensed with an external circuit.

For good sensitivity, the absorber should have a low heat capacity and good absorptivity over a large spectral range. Any supporting substrate must have a low heat capacity and a high thermal conductivity.

## Semiconductor Bolometers (cont.)

The resistor material is chosen to have low heat capacity and a high **thermal coefficient of resistance (TCR)**:

$$\alpha(T) = \frac{1}{R} \left( \frac{dR}{dT} \right)$$

For doped semiconductors at low temperatures,  $\alpha$  is large and negative, and  $R$  is low enough for coupling to low-noise amplifiers. (The TCR of metals is generally low.)

Thin metallic films with a sheet resistance matched to that of free space are often used as the absorbing element in bolometers. Alternatively, polarization-sensitive bolometers have been made by using two closely spaced orthogonal grids, each attached to its own thermistor.

For a constant current bias  $I$  and an input THz signal modulated at  $\omega$ , the responsivity of a bolometer is given by

$$R_V = \frac{\alpha I R}{G_e (1 + i\omega C/G_e)}; \quad G_e = G - \alpha I^2 R$$

Si bolometers typically achieve responsivities of  $\sim 10^7$  V/W and NEPs of  $\sim 10^{-14}$  W/Hz<sup>1/2</sup>. The response time of the bolometer depends on the heat capacity and thermal conductance  $\tau = C/G_e$ .

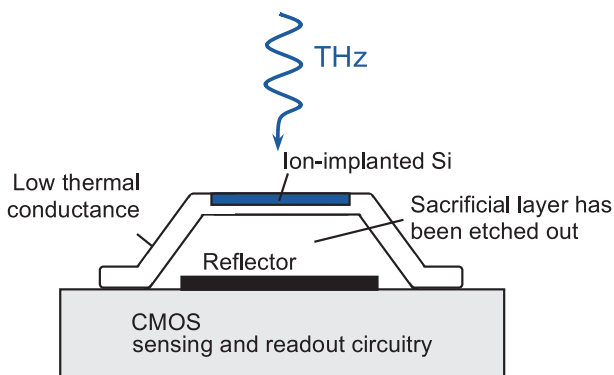
Since the heat capacity of the absorber scales with area, detection sensitivity is maximized for small absorber areas. Absorption efficiency can be increased by mounting the device in an integrating cavity and using **multimode horns** (e.g., Winston cones) to improve coupling to the incident THz radiation.

Random fluctuations in the flow of heat from the bolometer result in **phonon noise** ( $NEP \approx \sqrt{4k_B T G}$  for an approximately isothermal system), and most bolometers also have excess electrical noise at low frequencies (**1/f noise**).

## Microbolometer Arrays

Uncooled arrays of bolometers (**microbolometers**) are being investigated for imaging applications that require both high yields and moderate costs. These operate on the same principle as cooled semiconductor bolometers, where incoming THz radiation heats a sensor material and changes its resistance. The technology was initially developed for IR radiation but more recently has been optimized for THz.

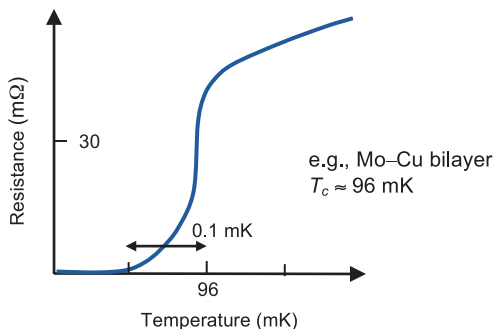
Microbolometers typically take advantage of standard Si processing technology; a substrate membrane supported by narrow legs is micromachined from an Si wafer using lithographic techniques. Ion implantation can be used to create the thermal sensing resistor in the substrate itself, and the whole process made compatible with CMOS fabrication techniques and materials. A reflecting layer placed  $\lambda/4$  below the sensor increases absorption efficiency.



As technology improves, large-format arrays of more-novel micromachined detectors are being developed, and the term *microbolometer* no longer refers to only the uncooled devices.

## Transition-Edge Sensors

Superconducting **transition-edge sensors (TESs)** consist of thin superconducting films biased to operate in the narrow temperature range, where they transition from the normal to superconducting (zero electrical resistance) state. Because the phase transition can be extremely sharp, a small increase in temperature dramatically changes the film resistance, making these devices useful as sensitive bolometers.



The superconducting **transition temperature**  $T_c$ , a parameter that is sensitive to processing conditions, is a critical factor in device performance. Most TESs are based on low- $T_c$  superconductors and are designed to have a transition temperature close to, e.g., 100 mK for dilution fridge operation or 400 mK if a  $^3\text{He}$  cryostat is used. Transition temperatures in this range can be achieved using techniques such as magnetic doping or, most commonly, the use of proximity-coupled superconductor-metal **bilayers**. Once biased with a constant voltage, incoming radiation heats the bolometer and causes its resistance to increase. This increase in turn reduces the power dissipated in the device and consequently lowers its temperature. Such a feedback mechanism can keep the operating temperature of TESs almost constant.

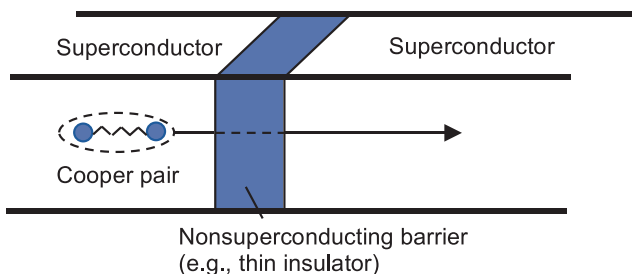
Readouts from an array of voltage-biased TES devices (which have low impedance) can be multiplexed using **SQUID** current amplifiers, so this is a promising technology for large-format low-noise arrays. NEPs of  $10^{-17} \text{ W/Hz}^{1/2}$  have already been achieved.



## SQUIDS

A **SQUID (superconducting quantum interference device)** is a sensitive magnetometer that can be used to measure extremely weak magnetic fields. The SQUID can be transformed into an amplifier for electrical currents by means of a coupling coil. More recently, SQUID multiplexers have made it possible to build large-format arrays of THz detectors.

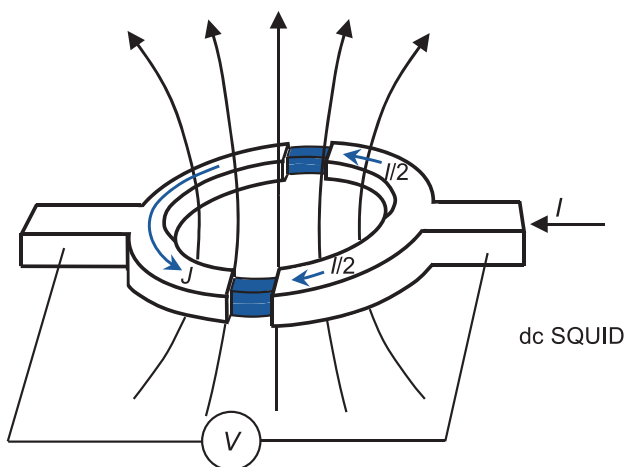
The operation of a SQUID is based on the phenomena of flux quantization in superconducting loops and the **Josephson effect**.



The Josephson effect refers to the ability of an electric current (Cooper-paired electrons) to flow, without any resistance, between two superconductors separated by a thin nonsuperconducting layer (a Josephson junction) through quantum tunneling. The maximum current that a Josephson junction can support, without developing any voltage across it, is known as its critical current  $I_c$ .

There are two main types of SQUIDS: ac or RF SQUIDS and the more sensitive dc SQUIDS. The dc SQUID consists of two Josephson junctions (the weak links) in parallel in a closed loop of superconductor.

### SQUIDS (cont.)



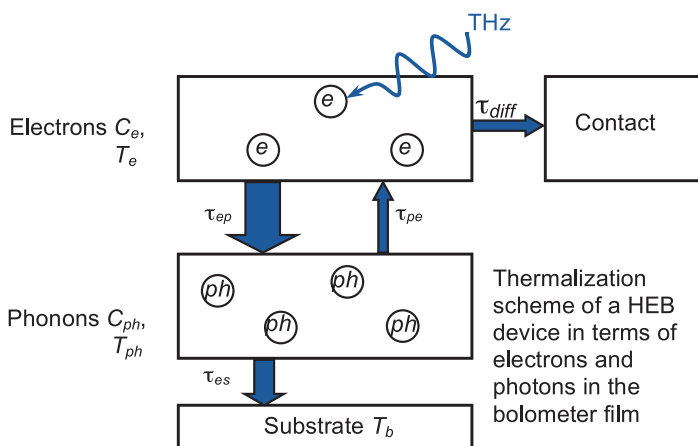
A dc input current  $I$  splits into two, and a current  $I/2$  flows through each of the two junctions. If a small external magnetic flux is applied to the superconducting loop, it responds with a screening current  $J$  to cancel this applied flux. The screening current flowing around the SQUID loop leads to a reduction in the critical current of the device from  $2I_c$  to  $(2I_c - 2J)$ .

Since the flux enclosed by the superconducting loop must be an integral number of flux quanta  $\Phi_0$  ( $2 \times 10^{-15}$  Wb), once the external flux is increased above  $\Phi_0/2$  it is energetically favorable for the SQUID to increase the enclosed flux to  $\Phi_0$ . Thus, the screening current changes direction every time the external flux increases by a half integer multiple of  $\Phi_0$ . If the SQUID is biased with a current slightly larger than  $2I_c$ , then the SQUID operates in the resistive mode, and the voltage is a periodic function of the applied magnetic field, and the period is equal to  $\Phi_0$ , which is a very small magnetic flux.

The SQUID can measure any physical quantity that can be transformed into magnetic flux, such as electrical current or voltage. The low noise, low power, and low impedance of these devices make them ideal amplifiers.

## Hot-Electron Bolometers

In a conventional bolometer, the crystal lattice absorbs energy and transfers it to free carriers, normally electrons, via phonons (collisions); thermal equilibrium between electrons and phonons is established instantly. In the case of **hot-electron bolometers (HEBs)**, the radiation is absorbed directly by electrons, which are only weakly coupled to the lattice. The term *hot electrons* refers to the non-equilibrium distribution of electrons, which can be described by the Fermi function but at a temperature that is elevated compared to that of the lattice. In semiconductors, the mobility of the carriers (resistance) depends on this temperature.



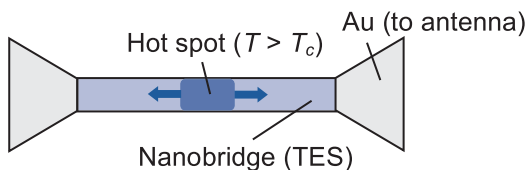
The electrons and lattice can be decoupled if either non-equilibrium phonons leave the detector quickly ( $\tau_{es} < \tau_{pe}$ ) or if the intensity of the THz radiation varies faster than  $1/\tau_{pe}$ . The first HEBs produced used n-type InSb crystal.

The electron temperature relaxes back to global equilibrium through the emission of phonons ( $\tau_{ep} \approx 10^{-7}$  s). The short relaxation time means that fast semiconductor hot-electron detectors can be designed that achieve NEPs of  $\sim 10^{-13}$  W/Hz<sup>1/2</sup>.

## Hot-Electron Bolometer Mixers

**Superconducting hot-electron bolometers (HEBs)** are similar in principle to TESs. In the case of HEBs, a fast response time is achieved by allowing THz radiation to be absorbed directly by electrons in the superconductor. The extremely rapid operation of superconducting HEBs has led to their use as mixers in heterodyne receivers.

The basic geometry of an HEB mixer consists of a thin-film superconducting nanobridge with thick normal-metal contacts at each end. DC biasing and irradiation with LO power creates an electron temperature profile across the bridge. The region where the temperature exceeds the critical temperature  $T_c$  is called the **hot spot**. Mixing occurs when incident THz radiation modulates the length of the resistive region, and therefore the device resistance, at the IF frequency.

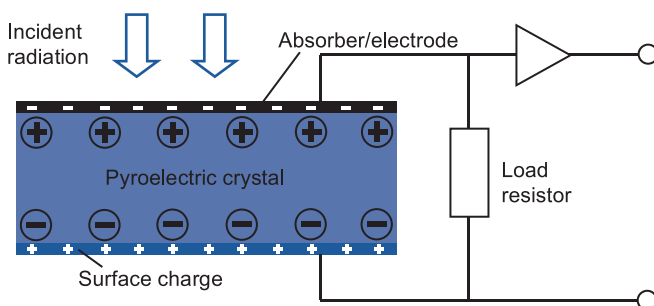


The thermal conductance for hot electrons must be high in order to achieve a fast response (to follow the IF). HEB mixers are classified as either **phonon cooled** or **diffusion cooled**, depending on the mechanism used to remove the thermal energy of the hot electrons. In a phonon-cooled mixer, energy is removed by electron–phonon collisions and subsequent escape of the phonons to the substrate.  $\tau_{es}$  (which is proportional to film thickness) must be shorter than the phonon–electron interaction time  $\tau_{pe}$ . Diffusion-cooled mixers, on the other hand, make use of the rapid diffusion of hot electrons from the ends of the nanobridge into the normal-metal contacts. The nanobridge must be short in order for this to be the dominant cooling mechanism. Typically, phonon-cooled mixers are made from ultrathin films of NbN, whereas diffusion-cooled devices use Nb or Al. HEBs have been successfully used as mixers at frequencies above 1 THz.

## Pyroelectric Detectors

**Pyroelectric detectors** are uncooled thermal detectors that exploit the variation with temperature of the spontaneous electric polarization in a pyroelectric (or **polar**) crystal.

Many crystals are pyroelectric, meaning that they possess a permanent electric dipole moment along one axis. The resulting charge expected on a surface that is cut perpendicular to this axis is normally neutralized by free carriers and is not observed in the steady state. However, if the temperature of the crystal changes, e.g., by absorbing incident THz radiation, the lattice spacing, and therefore the spontaneous polarization and surface charge, also change.



In a typical pyroelectrical detector, electrodes are placed on opposite faces of the crystal to form a capacitor, and current flows through the circuit to compensate for the change in charge. Pyroelectric detectors must therefore be used in ac mode by modulating the incident THz radiation.

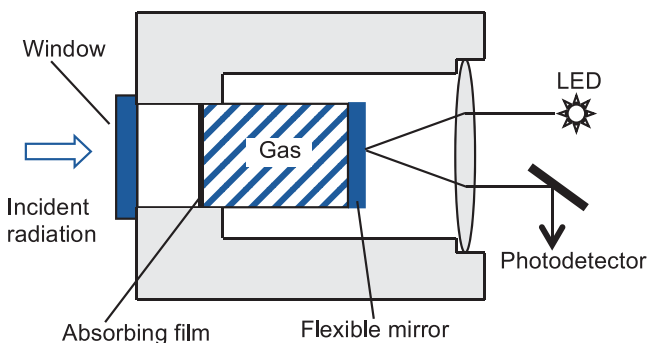
Commonly used pyroelectric crystals include triglycine sulphate (TGS), deuterated triglycine sulphate (DTGS), lithium tantalite ( $\text{LiTaO}_3$ ), and barium titanate ( $\text{BaTiO}_3$ ).

The responsivity of pyroelectric detectors can reach 1 kV/W for large load resistances with NEPs of  $\sim 10^{-9} \text{ W/Hz}^{1/2}$  ( $\sim 10$ -Hz modulation frequency).

## Golay Cells

**Golay cells** are uncooled thermal (incoherent) detectors that exploit the change in gas pressure with temperature.

A Golay cell consists of a semi-transparent absorbing film placed inside a pneumatic cell. A THz signal that is incident upon the absorber (blackened thin metal film on a substrate) heats the gas inside the cell. The gas then expands, and the pressure in the cell increases. The resulting increase in pressure deforms a flexible mirror fixed to the back of the cell, and this motion is detected by an optical reflectivity measurement. Modulated THz radiation at the input therefore appears as a modulated optical signal at the output.



Xenon is a gas commonly used in Golay cells because of its low thermal conductivity. The window and pneumatic chamber are also made of good heat insulators. The spectral range of the device is often determined by the transmission of the input window (20–1000  $\mu\text{m}$  for polyethylene).

Golay cells are sensitive room-temperature detectors but have a relatively slow response time. For modulation frequencies of a few tens of Hz, they have responsivities on the order of  $\text{kV/W}$  and NEPs of  $\sim 10^{-10} \text{ W/Hz}^{1/2}$ .

## Gaussian Beam Propagation

---

Consider one component,  $E(x, y, z)$ , of the complex electric field of a monochromatic, spatially coherent beam propagating in the positive  $z$  direction. It satisfies the time-independent, or **Helmholtz, wave equation**:

$$\nabla^2 E + k^2 E = 0, \quad \text{where } k = 2\pi/\lambda$$

If we rewrite the electric field component as

$$E(x, y, z) = u(x, y, z) \exp(-ikz)$$

then the Helmholtz equation becomes  $\nabla^2 u - 2ik\partial u/\partial z = 0$ .

If the beam is paraxial (largely collimated with an angular divergence less than about 30 deg), then  $u(x, y, z)$  varies only slowly with  $z$  over a distance comparable with a wavelength, and  $|\partial^2 u/\partial z^2| \ll |k\partial u/\partial z|$ . In this case, the Helmholtz equation becomes the **paraxial wave equation**

$$\partial^2 u/\partial x^2 + \partial^2 u/\partial y^2 - 2ik\partial u/\partial z = 0$$

The **Gaussian beam modes** that play such an important role in long-wavelength optics are the solutions to this equation. They are the **Hermite polynomials** in rectangular coordinates, and the **Laguerre polynomials** in cylindrical coordinates, both having a Gaussian envelope function. The form of the amplitude distribution of these modes remains unchanged as they propagate. Since they constitute complete orthonormal sets, any arbitrary solution of the paraxial wave equation can be expressed as a superposition of Gaussian modes. Typically, the lowest-order modes are dominant.

This is a scalar treatment, and  $E$  can be thought of as representing the dominant (e.g., co-polar) component of an electrical field. If detailed information about the polarization properties of a beam are required, then a different, vector approach is necessary.

## Complex Radius of Curvature

A simple solution to the paraxial wave equation can be written in the form

$$u(x, y, z) = \frac{u_0}{(q_0 + z)} \exp\left(-i \frac{k(x^2 + y^2)}{2(q_0 + z)}\right)$$

If we choose  $q_0 = ikw_0^2/2$ , then at  $z = 0$  we obtain a Gaussian-shaped beam:

$$u(x, y, z) \propto \exp\left(-\frac{(x^2 + y^2)}{w_0^2}\right)$$

with a planar phase front, and  $1/e$  in amplitude beam radius of  $w_0$ . On propagating away from  $z = 0$ , the exponent in the expression for  $u(x, y, z)$  has both real and imaginary terms:

$$\begin{aligned} u(x, y, z) &= \frac{u_0}{q(z)} \exp\left(-i \frac{k(x^2 + y^2)}{2(q(z))}\right) \\ &= \frac{u_0}{q(z)} \exp\left[\left(-i \frac{k(x^2 + y^2)}{2q_r(z)}\right) - \left(\frac{k(x^2 + y^2)}{2q_i(z)}\right)\right] \end{aligned}$$

where  $q(z) = q_0 + z$ , and  $1/q(z) = 1/q_r(z) - i(1/q_i(z))$ .

The exponent has both an imaginary and a real quadratic transverse variation. We can associate the imaginary part with the phase of a spherical wave front—with real radius of curvature  $R(z)$ —in the paraxial limit, and the real part with a Gaussian amplitude profile of radius  $w(z)$ :

$$u(x, y, z) = \frac{u_0}{q(z)} \exp\left[\left(-i \frac{k(x^2 + y^2)}{2R(z)}\right) - \left(\frac{(x^2 + y^2)}{w^2(z)}\right)\right]$$

so

$$R(z) = \left(\operatorname{Re}\left[\frac{1}{q(z)}\right]\right)^{-1} \quad \text{and} \quad w(z) = \sqrt{\frac{-\lambda}{\pi \operatorname{Im}\left[1/q(z)\right]}}$$

$q(z)$  is known as the **complex radius of curvature**, complex beam parameter, or Gaussian beam parameter.



## Gaussian Beam Parameters

The complex beam parameter in the simple solution to the paraxial wave equation is given by

$$\frac{1}{q(z)} = \frac{1}{R(z)} - i \frac{\lambda}{\pi w^2(z)}$$

where  $w(z)$  is the Gaussian **beam radius**,  $R(z)$  is the Gaussian beam **phase radius of curvature**, and

$$q(z) = i \frac{\pi w_0^2}{\lambda} + z$$

From these we can derive the variation of beam radius and radius of curvature with propagation distance  $z$ :

$$w(z) = w_0 \left[ 1 + \left( \frac{\lambda z}{\pi w_0^2} \right)^2 \right]^{1/2}$$

$$R(z) = z + \frac{1}{z} \left( \frac{\pi w_0^2}{\lambda} \right)^2$$

The solution to the wave equation

$$u(x, y, z) = \frac{u_0}{q(z)} \exp \left[ \left( -i \frac{k(x^2 + y^2)}{2R(z)} \right) - \left( \frac{x^2 + y^2}{w^2(z)} \right) \right]$$

can be re-expressed in the usual form:

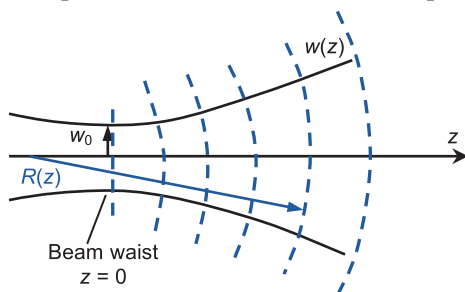
$$\frac{u_0}{q_0} \frac{w_0}{w(z)} \exp \left[ \left( -i \frac{k(x^2 + y^2)}{2R(z)} \right) - \left( \frac{x^2 + y^2}{w^2(z)} \right) + i \phi_0(z) \right]$$

where  $\phi_0(z)$ , known as the Gaussian beam **phase slippage**, is given by

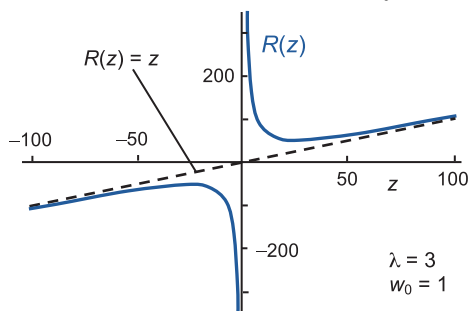
$$\phi_0(z) = \tan^{-1} \left( \frac{\lambda z}{\pi w_0^2} \right)$$

## Gaussian Beam Parameters (cont.)

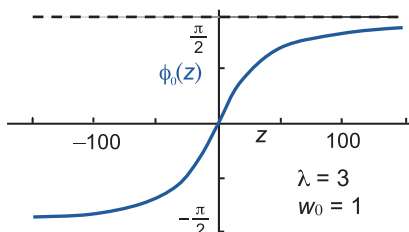
At  $z = 0$ , a position known as the **beam waist**, the solution to the paraxial wave equation is a fundamental Gaussian beam of radius  $w(0) = w_0$  and infinite phase radius of curvature  $R(0) = \infty$ .  $w_0$  is known as the **beam waist radius**, and it is the minimum value of  $w(z)$ . As the beam propagates it spreads out, and its on-axis amplitude drops.



Away from the beam waist, the phase radius of curvature first drops and then increases. Eventually,  $R(z) \rightarrow z$ .



$\phi_0(z)$  represents the phase slippage of the Gaussian mode with respect to a plane wave  $E = E_0 \exp(-ikz)$ . In the far field,  $\phi_0(z) \rightarrow \pi/2$ .



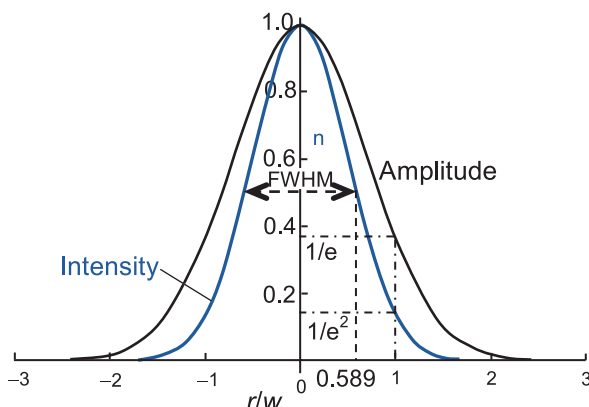
## Beamwidth

The electric field amplitude of a Gaussian beam can be written as

$$E(r, z) = E_0 \exp\left(-\frac{r^2}{w^2(z)}\right)$$

with  $r$  being the off-axis distance,  $E_0$  the on-axis amplitude, and  $w$  the Gaussian beam radius that varies with propagation distance  $z$ . The intensity is

$$I(r, z) = I_0 \exp\left(-\frac{2r^2}{w^2(z)}\right)$$



The beam radius  $w$  is the off-axis distance at which the amplitude falls to  $1/e$  of its on-axis value and the intensity to  $1/e^2$  of its on-axis value.

The fraction of total power contained within a radius  $r$  is

$$\frac{P(r)}{P_\infty} = 1 - \exp\left(-\frac{2r^2}{w^2(z)}\right)$$

The **full-width-half-maximum** (FWHM) or **half-power beamwidth** (HPBW) is the full width of the beam at half its maximum intensity. For a Gaussian beam,

$$FWHM = \sqrt{2 \ln(2)} w$$

## Edge Taper

The ratio of power at the edge of an aperture to that on-axis is known as the **edge taper**:

$$T_e = \frac{P(a)}{P(0)}$$

where  $a$  is the aperture radius. Expressed in decibels,

$$T_e(\text{dB}) = -10\log_{10}(T_e)$$

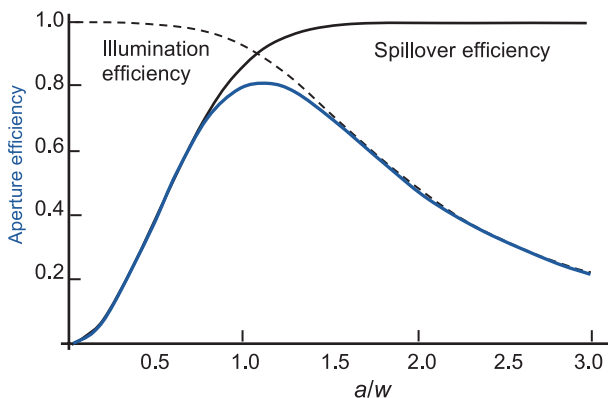
Under illumination by a Gaussian beam of radius  $w$ ,

$$T_e(\text{dB}) = \frac{20}{\ln(10)} \left( \frac{a}{w} \right)^2 \approx 8.686 \left( \frac{a}{w} \right)^2$$

Because the illumination decreases toward the edges of an aperture, its effective area can be considered less than its physical area. The **illumination** (or taper) **efficiency** is a measure of the nonuniformity of a beam over an aperture (the power coupling coefficient between the beam and a plane wave).

The **spillover efficiency** is a measure of the percentage of power in the beam intercepted by an aperture.

The maximum **aperture efficiency** (product of illumination and spillover efficiency) is 81%, which occurs for an aperture radius  $a = 1.12w$  or for an edge taper  $T_e = 10.9$  dB.



## Truncation and Spillover (Gaussian Beams)

**Truncation** leads to power loss and produces additional diffraction effects on the subsequent propagation of a beam. Determining the acceptable level for such effects informs the minimum size of components in an optical system.

For symmetric truncation of a pure Gaussian beam, the fraction of the power lost is a function of the ratio of the truncation radius to the beamwidth  $r_t/w$ :

$$K_{lost}(r_t/w) = 1 - \exp(-2r_t^2/w^2) = 1 - T_e(r_t/w)$$

where  $T_e(r_t/w)$  is the edge taper. The effective beamwidth is also reduced at the truncation plane. This reduction can have the effect of increasing or decreasing the corresponding waist size for the beam, depending on the phase slippage of the beam up to that point.

As a general rule, the minimum aperture radius is often set to  $r_t = 2w$ ; however, for non-Gaussian beams this can lead to losses greater than 1%, so for a safer margin it is better to set  $r_t = 4w$ .

At a reflecting or refracting component, the truncated power—if not terminated—results in **spillover** at the sides of the component. In an antenna receiver or transmitter system, this can result in power being received from or transmitted in unexpected directions, possibly degrading the system performance.

In the case of a Gaussian beam, there is a simple relationship between spillover power and edge taper. The fraction of the spillover power is:

$$K_{spillover}(r_t/w) = \exp(-2r_t^2/w^2) = T_e(r_t/w)$$

## Truncation and Spillover (Non-Gaussian Beams)

Because typical beams are not only tapered (as a Gaussian beam), but often also possess, for example, a sidelobe structure, the precise analysis of truncation effects is important.

Truncation effects on non-Gaussian beams in a beam guide can be treated more rigorously using **Gaussian beam mode analysis (GBMA)** and introducing **scattering matrices**. If axis-symmetric truncation takes place, this leads to simple analytical forms for the truncation.

For non-Gaussian beams, the spillover efficiency is defined as

$$\eta_{spillover} = \frac{\iint_{r_t} |E|^2 dS}{\iint_{\infty} |E|^2 dS}$$

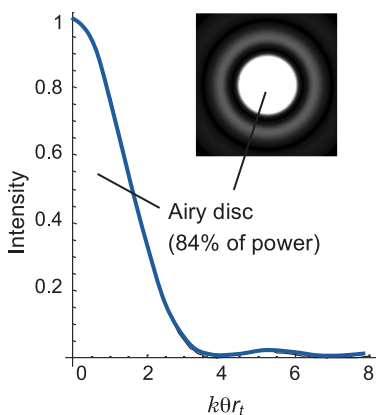
i.e., the fraction of the incident power that is transmitted (or blocked) by the truncating aperture (or blockage) of radius  $r_t$ .

In addition to power loss, a truncating aperture can have a significant effect on the shape of a propagating beam. In the case of uniform illumination of a circular aperture, the far-field pattern takes on the familiar **Airy pattern** intensity profile.

The intensity falls to zero at an off-axis angle of

$$\theta = 1.22\lambda/D_t,$$

where  $D_t = 2r_t$ , the diameter of the truncating aperture. This corresponds to the **Rayleigh criterion** for resolving two nearby point sources.

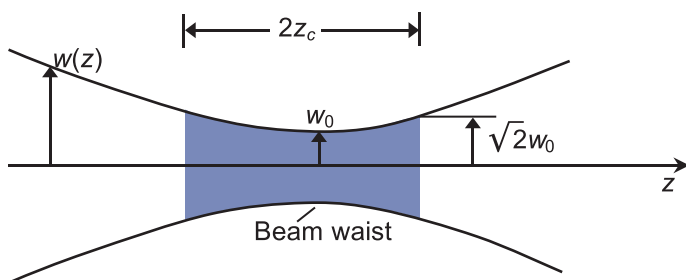


## Confocal Distance

The **confocal distance** (or **Rayleigh range**)  $z_c$  of a Gaussian beam of waist radius  $w_0$  is defined as

$$z_c = \frac{\pi w_0^2}{\lambda}$$

A Gaussian beam remains quasi-collimated ( $w$  increases by a factor less than  $\sqrt{2}$ ) within a distance  $z_c$  of the waist, so  $z_c$  represents a kind of “depth of focus.”



Using this definition for the confocal distance, the Gaussian beam parameters can be written as

$$w(z) = w_0 \left[ 1 + \left( \frac{z}{z_c} \right)^2 \right]^{1/2}$$

$$R(z) = z + \frac{z_c^2}{z}$$

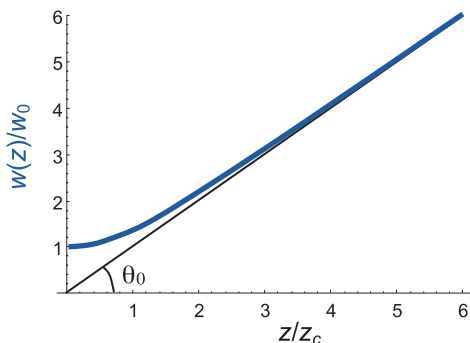
$$\phi_0(z) = \tan^{-1} \left( \frac{z}{z_c} \right)$$

and the region around the beam can be divided into a near ( $z \ll z_c$ ) and a far field ( $z \gg z_c$ ).

The confocal distance is sometimes denoted as  $z_0$ ,  $z_R$ , or  $\hat{z}$ .

## Far-Field Divergence

A Gaussian beam remains essentially collimated within the confocal distance but increases monotonically thereafter.



In the far field ( $z \gg z_c$ ), the beam radius grows linearly with  $z$ .

If the beam radius is defined in terms of the angle it makes with the propagation axis

$$\theta = \tan^{-1} \left( \frac{w}{z} \right)$$

then in the far field we obtain the **asymptotic beam growth angle**  $\theta_0$  given by

$$\theta_0 = \tan^{-1} \left( \frac{\lambda}{\pi w_0} \right)$$

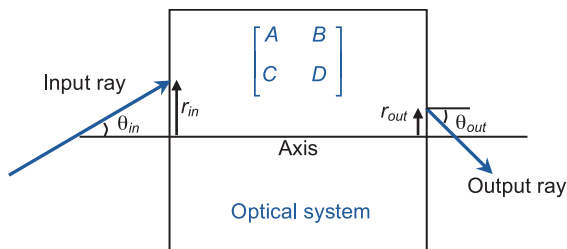
or in the paraxial limit,

$$\theta_0 \approx \frac{\lambda}{\pi w_0} \text{ radians}$$



## Ray Matrices and Gaussian Beam Transformation

For multi-element optical systems, we can use the convenient *ABCD* ray transfer matrix formalism that is commonly used in geometrical optics.



The ***ABCD*** or **ray matrix** for the optical element operates on an input ray displacement and slope. It transforms  $(r_{in}, \theta_{in})$  at the input plane to  $(r_{out}, \theta_{out})$  at the output plane according to the matrix equation

$$\begin{bmatrix} r_{out} \\ \theta_{out} \end{bmatrix} = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} r_{in} \\ \theta_{in} \end{bmatrix}$$

For Gaussian beams, the four parameters of the ray matrix operate on the complex radius of curvature  $q$ , giving

$$q_{out} = \frac{A \cdot q_{in} + B}{C \cdot q_{in} + D}$$

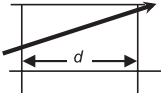
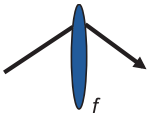
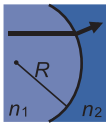

and on the phase slippage, giving

$$\phi_{out} = \phi_{in} - \arg \left[ A + B \left( 1/q_{in} \right) \right]$$

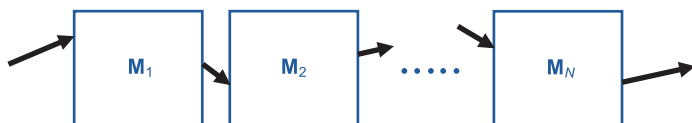
The beam radius and phase radius of curvature can be recovered from  $q$ .

## Ray Matrices and Gaussian Beam Transformation (cont.)

Common ray transformation matrices:

Element	Diagram	Matrix
Propagation a distance $d$ in free space or in a medium of constant refractive index		$\begin{bmatrix} 1 & d \\ 0 & 1 \end{bmatrix}$
Thin lens of focal length $f$ ( $f > 0$ for converging)		$\begin{bmatrix} 1 & 0 \\ -\frac{1}{f} & 1 \end{bmatrix}$
Refraction at a curved dielectric interface (normal incidence, $R > 0$ for center of curvature before interface)		$\begin{bmatrix} 1 & 0 \\ \frac{(n_2 - n_1)}{R n_2} & \frac{n_1}{n_2} \end{bmatrix}$
Reflection from a curved mirror (normal incidence, $R > 0$ for center of curvature before interface)		$\begin{bmatrix} 1 & 0 \\ -\frac{2}{R} & 1 \end{bmatrix}$

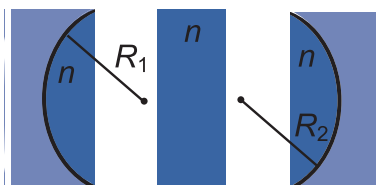
# Ray Matrices and Gaussian Beam Transformation (cont.)



For a multi-element system made up of a series of optical components, the full  $ABCD$  matrix is given by the product of the corresponding  $ABCD$  matrices for the individual components, including any distance propagated in free space or in a medium. The matrices must be cascaded in reverse order so that the first optical element is on the right-hand side of the product.

$$\mathbf{M}_{total} = \mathbf{M}_n \cdot \mathbf{M}_{n-1} \cdots \mathbf{M}_2 \cdot \mathbf{M}_1 = \begin{bmatrix} A & B \\ C & D \end{bmatrix}_{total}$$

For example, the  $ABCD$  matrix for a lens of thickness  $d$ , with spherical surfaces of radii  $R_1$  and  $R_2$ , made of a material of refractive index  $n$ , in air ( $n = 1$ ), is



$$\begin{aligned} \mathbf{M} &= \begin{bmatrix} 1 & 0 \\ \frac{1-n}{R_2} & n \end{bmatrix} \cdot \begin{bmatrix} 1 & d \\ 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 \\ \frac{n-1}{R_1 n} & \frac{1}{n} \end{bmatrix} \\ &= \begin{bmatrix} 1 + \frac{(n-1)d}{nR_1} & \frac{d}{n} \\ -(n-1)\left(\frac{1}{R_2} - \frac{1}{R_1} + \frac{(n-1)d}{nR_1 R_2}\right) & 1 - \frac{(n-1)d}{nR_2} \end{bmatrix} \\ &\quad (R_1 \text{ is negative and } R_2 \text{ is positive.}) \end{aligned}$$

## Higher-Order Modes (Cylindrical Coordinates)

A Gaussian beam is the simplest solution to the paraxial wave equation and suffices in many cases to describe beam propagation in an optical system. There are situations, however, when we need to model more complex field distributions, and in these cases we can use **higher-order beam-mode** solutions.

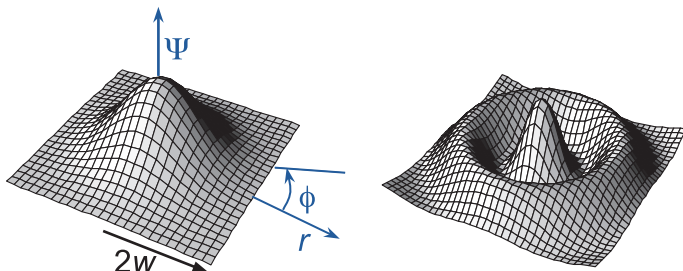
Higher-order modes are characterized by the same beam radius  $w(z)$  and phase radius of curvature  $R(z)$  as the fundamental mode; only the phase slippage differs. The most appropriate mode set to use depends on the symmetry of the system being modeled. In cylindrical coordinates we can use **Gaussian-Laguerre** modes:

$$\begin{aligned}\Psi_{pm}(r, \phi, z) = & \left[ \frac{2p!}{\pi(p+m)!} \right]^{0.5} \frac{1}{w(z)} \left[ \frac{\sqrt{2}r}{w(z)} \right]^m L_p^m \left( \frac{2r^2}{w^2(z)} \right) \\ & \cdot \exp \left[ \frac{-r^2}{w^2(z)} \right] \exp[-ikz] \exp \left[ -i \frac{\pi r^2}{\lambda R(z)} \right] \\ & \cdot \exp[i(2p+m+1)\phi_0(z)] \exp(im\phi)\end{aligned}$$

where  $L_p^m$  are generalized Laguerre polynomials. The phase slippage is now

$$\phi(z) = (2p + m + 1)\phi_0(z)$$

where  $\phi_0(z)$  is the phase slippage of the fundamental mode as before. The figure shows field distributions of the  $\Psi_{00}$  (left) and  $\Psi_{30}$  (right) Gaussian-Laguerre modes at a waist location.



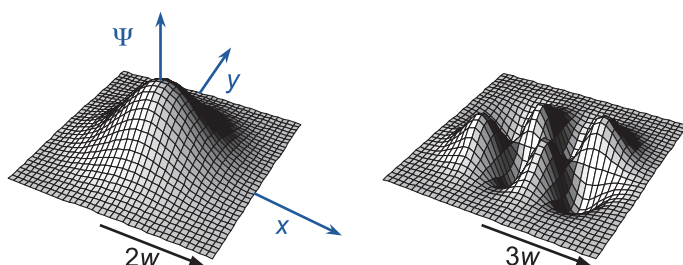
## Higher-Order Modes (Rectangular Coordinates)

In rectangular coordinates we use **Gaussian–Hermite** modes:

$$\begin{aligned}\Psi_{mn}(x, y, z) = & \left[ \frac{1}{\pi w_x(z) w_y(z) 2^{m+n-1} m! n!} \right]^{0.5} \\ & \cdot H_m \left[ \frac{\sqrt{2}x}{w_x(z)} \right] H_n \left[ \frac{\sqrt{2}y}{w_y(z)} \right] \cdot \exp \left[ -\frac{x^2}{w_x^2(z)} - \frac{y^2}{w_y^2(z)} \right] \\ & \cdot \exp[-ikz] \cdot \exp \left[ -i \frac{\pi x^2}{\lambda R_x(z)} - i \frac{\pi y^2}{\lambda R_y(z)} \right] \\ & \cdot \exp \left[ i \frac{(2m+1)}{2} \phi_{0x}(z) + i \frac{(2n+1)}{2} \phi_{0y}(z) \right]\end{aligned}$$

where  $H_m$  are Hermite polynomials, and the modes can have different widths and phase radii in the  $x$  and  $y$  directions. Again, there is a mode-dependent phase-slippage term, but  $w(z)$  and  $R(z)$  are given by the same expressions as are used for the fundamental mode.

The figure shows field distributions of the  $\Psi_{00}$  (left) and  $\Psi_{31}$  (right) Gaussian–Hermite modes at a waist location. The higher-order Gaussian–Hermite and Gaussian–Laguerre modes consist of polynomials superimposed on the fundamental Gaussian mode.



## Mode Coefficients

Gaussian–Hermite and Gaussian–Laguerre modes constitute complete orthonormal sets of modes that are each solutions to the paraxial wave equation. Any arbitrary solution of this wave equation can therefore be expressed as a superposition of Gaussian modes, the particular mode set,  $\Psi_i$ , being chosen as appropriate for the symmetry of the problem:

$$E(\mathbf{r}) = E(x, y, z) = \sum_i A_i \Psi_i(x, y, z; w(z), R(z), \phi_i(z))$$

where  $E(\mathbf{r})$  is the paraxial field,  $A_i$  are the **mode coefficients**,  $w(z)$ ,  $R(z)$ , and  $\phi_i(z)$  are the beam radius, phase radius, and phase slippage, respectively, as before. If the mode coefficients are known, they can be used to resynthesize the beam at any plane in an optical system.

If the field is known over a surface  $S$ , then the mode coefficients are determined by calculating the overlap integrals

$$A_i = \iint_S E(x, y, z) \cdot \Psi_i^*(x, y, z; w(z), R(z), \phi_i(z)) dS$$

Because of the limited spatial-frequency content of many long-wavelength systems, if the mode set that maximizes the power in the fundamental mode is chosen, a source field can very often be represented with high accuracy by the sum of only a few modes.

If optical components such as mirrors and lenses are treated as perfect phase transformers with no truncation, then the mode coefficients remain the same as the beam propagates.  $w(z)$  and  $R(z)$  vary with propagation distance in the same way for all modes; the phase-slippage term  $\phi_i(z)$  is mode-dependent. It is the change in a mode's phase-slippage term, with respect to other modes, that results in a beam pattern that evolves with propagation distance.

Mode Coefficients (cont.)

Example of Gaussian beam mode propagation:

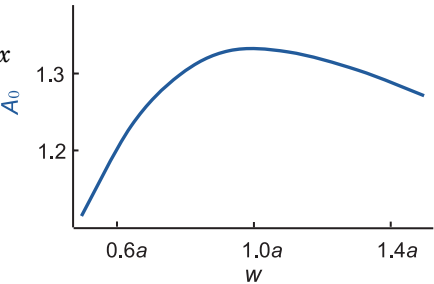
Consider (in one dimension) a plane wave propagating from an aperture of length  $2a$ , modeled using Gaussian–Hermite modes

$$\Psi_m(x,z=0;w,R=\infty)=\left[\frac{\sqrt{2}}{\sqrt{\pi}w2^mm!}\right]^{0.5}\cdot H_m\left(\frac{\sqrt{2}x}{w}\right)\cdot \exp\left[-\frac{x^2}{w^2}\right]$$

(neglecting the overall plane-wave phase factor, and setting  $R=\infty$  to match the flat phase of the input wave). An efficient choice of mode set maximizes the power in the fundamental mode  $A_0^2$ .  $A_0$  is calculated from the overlap integral

$$A_0=\int_{-a}^a\Psi_m(x,z;w,R)\,dx$$

and is maximized by setting  $w=1.010a$ .



With this choice of  $w$  and  $R$ , the mode coefficients of the first ten modes are:

Mode number	Mode coefficient
0	1.334
1	0
2	0.000
3	0
4	−0.251
5	0
6	0.211
7	0
8	−0.096
9	0

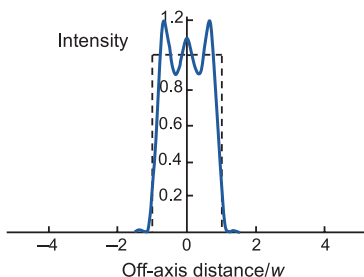
## Mode Coefficients (cont.)

The field at the aperture is reconstructed using

$$E(x, z=0) = \sum_{m=0}^M A_m \Psi_m(x, 0; w, \infty)$$

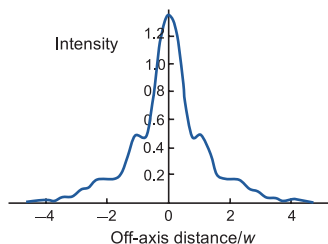
To reconstruct the field at any other plane, the Gaussian beam parameters  $w(z)$ ,  $R(z)$ , and  $\phi_0(z)$  are first calculated, and

$$E(x, z) = \sum_{m=0}^M A_m \Psi_m(x, z; w(z), R(z)) \exp \left[ i \left( m + \frac{1}{2} \right) \phi_0(z) \right]$$



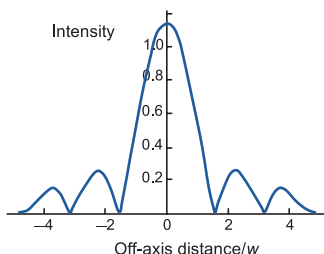
$$\begin{aligned} a &= 1 \text{ mm} & w &= 1.01 \text{ mm} \\ \lambda &= 1 \text{ mm} & R &= \infty \\ z &= 0 \text{ mm} \end{aligned}$$

$$\begin{aligned} \phi_0 &= 0 \\ F \text{ (Fresnel no.)} &= a^2/z\lambda \\ &= \infty \end{aligned}$$



$$\begin{aligned} a &= 1 \text{ mm} & w &= 1.19 \text{ mm} \\ \lambda &= 1 \text{ mm} & R &= 7.14 \text{ mm} \\ z &= 2 \text{ mm} \end{aligned}$$

$$\begin{aligned} \phi_0 &= 31.5 \text{ deg} \\ F &= 0.5 \end{aligned}$$



$$\begin{aligned} a &= 1 \text{ mm} & w &= 63.03 \text{ mm} \\ \lambda &= 1 \text{ mm} & R &= 200.05 \text{ mm} \\ z &= 200 \text{ mm} \end{aligned}$$

$$\begin{aligned} \phi_0 &= 89.10 \text{ deg} \\ F &= 0.005 \end{aligned}$$



## Power Coupling Efficiency

The **coupling efficiency** (or fractional power coupling) between two beams,  $a$  and  $b$ , is defined as

$$K_{ab} = \frac{|\iint_S \psi_a^* \psi_b dS|^2}{\iint_S |\psi_a|^2 dS \cdot \iint_S |\psi_b|^2 dS}$$

where  $\psi_a$  and  $\psi_b$  are the relevant scalar fields over a plane transverse to the propagation axis of one of the fields. If we express the two beams as a modal sum with the same normalized Gaussian beam mode set (i.e.,  $\psi_a = \sum_n A_n \Psi_n$ ,  $\psi_b = \sum_n B_n \Psi_n$ , and  $\iint_\infty |\Psi_n|^2 dS = 1$ ), then the coupling efficiency can be re-expressed as

$$K_{ab} = \frac{\left| \sum_n A_n^* B_n \right|^2}{\sum_n |A_n|^2 \sum_n |B_n|^2}$$

If the beams are normalized a relative power coupling efficiency is obtained, where  $K_{ab} = \left| \sum_n A_n^* B_n \right|^2$ . For a vector model, the power coupling efficiency can be defined as

$$K_{ab} = \frac{|\iint_S \mathbf{E}_a \times \mathbf{H}_b^* dS|^2}{|\iint_S \mathbf{E}_a \times \mathbf{H}_a^* dS| \cdot |\iint_S \mathbf{E}_b \times \mathbf{H}_b^* dS|}$$

where  $\mathbf{E}$  and  $\mathbf{H}$  are electric and magnetic fields, respectively, for the two beams over a plane transverse to the propagation axis of one field. For paraxial beams, the orthogonal components of the two fields can be expressed as modal sums of the form  $\mathbf{E}_a = \sum_n (A_n^h \hat{\mathbf{e}}_h + A_n^v \hat{\mathbf{e}}_v) \Psi_n$ , and then we can re-express the coupling efficiency in terms of these modal sums as

$$K_{ab} = \left| \sum_n (A_n^{h*} B_n^h + A_n^{v*} B_n^v) \right|^2$$

assuming both fields are normalized so that

$$\sum_n \left( |A_n^h|^2 + |A_n^v|^2 \right) = 1$$

## Gaussicity

The **Gaussicity** of a beam is its power coupling to a best-fit pure Gaussian field over some plane transverse to the propagation direction. In terms of a scalar model, this can be defined for a field  $\psi_F$  as

$$K_G = \frac{\left| \int_A \psi_G^*(w, R) \psi_F dS \right|^2}{\int_A |\psi_F|^2 dS \cdot \int_\infty |\psi_G(w, R)|^2 dS}$$

where  $\psi_G$  is the best-fit Gaussian to the beam in question. The pure Gaussian is assumed to extend over the whole plane and hence is integrated to infinity. The Gaussian is the best fit to both the amplitude distribution and the phase curvature of the field. The integral can be undertaken at any convenient transverse plane because the power coupling is independent of the location of this plane.

For a vector model, the Gaussicity can be defined in terms of the coupling to a linearly polarized best-fit Gaussian field:

$$K_G(w, R) = \frac{\left| \int_A \mathbf{e}_G^*(w, R) \cdot \mathbf{e}_F dS \right|^2}{\int_A |\mathbf{e}_F|^2 dS \cdot \int_\infty |\mathbf{e}_G(w, R)|^2 dS}$$

where the best fit Gaussian  $\mathbf{e}_G$  is aligned with the dominant co-polar field of  $\mathbf{e}_F$ . A similar definition can be used for circularly polarized beams.

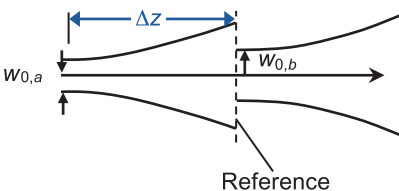
For example:

Horn type	Gaussicity
Corrugated	0.981
Conical	0.866
Rectangular	0.872
Diagonal	0.840
Ultra-Gaussian	0.999

## Mismatched Beams and Tolerancing

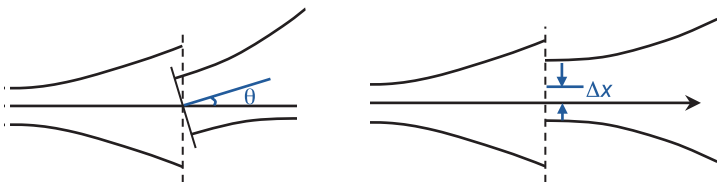
In analyzing losses in nonideal systems, the coupling of beams becomes an issue; for example, the coupling of an incident beam from a source to a beam that characterizes a component such as a detector is problematic. Losses can be due to mechanical **misalignment**, **defocusing**, or other **tolerancing** effects.

A defocused system occurs when two beams have their optical axes aligned, but the beam waists do not coincide. Thus, there is a mismatch in the beamwidth  $w$  and phase curvature  $R$ . In this case, the coupling efficiency of the two beams drops from unity to

$$K_{ab} = \frac{4}{\left(\frac{w_{0,b}}{w_{0,a}} + \frac{w_{0,a}}{w_{0,b}}\right)^2 + \left(\frac{\lambda \Delta z}{\pi w_{0,a} w_{0,b}}\right)^2}$$


assuming that we can approximate the beams as Gaussians with beam waist parameters of  $w_{0,a}$  and  $w_{0,b}$ , respectively, and  $\Delta z$  is the relative displacement of the two waists.

When one beam is tilted with respect to the other, the mismatch gives rise to a phase slope in one of the fields relative to the other across the plane where their optical axes intersect. For Gaussian beams that otherwise have the same beam parameters (and  $\Delta z = 0$ ), the fractional power coupling between the beams is given by  $K_{ab} = \exp\left(-(\pi w \theta / \lambda)^2\right)$ , where  $w$  is the beamwidth at the plane of intersection. For laterally shifted (by  $\Delta x$ ) but otherwise matched beams, the coupling efficiency is  $K_{ab} = \exp\left(-[\Delta x / w]^2\right)$ .



## Scattering Matrix Formulism

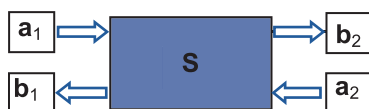
**Scattering matrices** describe the propagation of a paraxial beam through a nonperfect (**aberrating** or truncating, for example) optical system. In this case, the basis Gaussian beam modes  $\Psi_m$  become distorted and are no longer pure modes of propagation.

The transmission of the field from the (transverse) input plane at port 1 to the (transverse) output plane at port 2 can be expressed by the scattering equation

$$\Psi_n^{SC} = \sum_m S_{mn} \Psi_m,$$

i.e, the original scattered mode has now been decomposed into a sum of true modes at the output plane.

For any coherent incident beam  $E_{in}(1) = \sum_n a_{n1} \Psi_n(1)$  at port 1, the transmitted field at port 2 can be written as  $E_{out}(2) = \sum_n a_{n1} \Psi_n^{SC} = \sum_m b_{m2} \Psi_m(2)$ , where the new mode coefficients  $b_{m2}$  are given by  $b_{m2} = \sum_n S_{mn} a_{n1}$ . In matrix notation, this scattering of power between modes can be written as  $\mathbf{b}_2 = \mathbf{S} \cdot \mathbf{a}_1$ , where  $\mathbf{a}_1 = [a_{11}, a_{21}, a_{31} \dots]^T$ , etc.



This approach can be extended to include **partial reflections**, giving rise to backward-moving waves and **standing waves** in the system. In this case, at each scattering component we have both “incident” and “reflected” waves, represented by incident and reflected beam-mode coefficient vectors  $\mathbf{a}_1$  and  $\mathbf{b}_1$ , and  $\mathbf{a}_2$  and  $\mathbf{b}_2$  at both input and output ports, respectively.

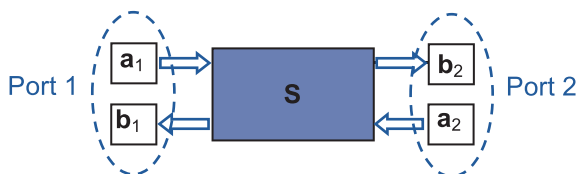
The full scattering matrix formulation (with 4 scattering matrices  $\mathbf{S}_{11}$ ,  $\mathbf{S}_{12}$ ,  $\mathbf{S}_{21}$ , and  $\mathbf{S}_{22}$ ) yields the matrix equation

$$\begin{pmatrix} \mathbf{b}_1 \\ \mathbf{b}_2 \end{pmatrix} = \begin{pmatrix} \mathbf{S}_{11} \cdot \mathbf{a}_1 + \mathbf{S}_{12} \cdot \mathbf{a}_2 \\ \mathbf{S}_{21} \cdot \mathbf{a}_1 + \mathbf{S}_{22} \cdot \mathbf{a}_2 \end{pmatrix} = \begin{pmatrix} \mathbf{S}_{11} & \mathbf{S}_{12} \\ \mathbf{S}_{21} & \mathbf{S}_{22} \end{pmatrix} \begin{pmatrix} \mathbf{a}_1 \\ \mathbf{a}_2 \end{pmatrix}$$

## Transmission Matrices

A **transmission matrix** that directly relates the fields at the output at port 2 to the input at port 1 of a mode-scattering system is expressed as

$$\begin{pmatrix} \mathbf{b}_2 \\ \mathbf{a}_2 \end{pmatrix} = \begin{pmatrix} \mathbf{T}_{11} \cdot \mathbf{a}_1 + \mathbf{T}_{12} \cdot \mathbf{b}_1 \\ \mathbf{T}_{21} \cdot \mathbf{a}_1 + \mathbf{T}_{22} \cdot \mathbf{b}_1 \end{pmatrix} = \begin{pmatrix} \mathbf{T}_{11} & \mathbf{T}_{12} \\ \mathbf{T}_{21} & \mathbf{T}_{22} \end{pmatrix} \begin{pmatrix} \mathbf{a}_1 \\ \mathbf{b}_1 \end{pmatrix}$$



Usually it is more straightforward to calculate the S-parameter scattering matrices  $\mathbf{S}_{11}$ ,  $\mathbf{S}_{12}$ ,  $\mathbf{S}_{21}$ , and  $\mathbf{S}_{22}$ , as they directly describe how an incident field at either port 1 or port 2 is divided between the corresponding resulting reflected and transmitted fields. The corresponding transmission matrices  $\mathbf{T}_{ij}$  can then be solved in terms of  $\mathbf{S}_{ij}$  to give:

$$\begin{pmatrix} \mathbf{T}_{11} & \mathbf{T}_{12} \\ \mathbf{T}_{21} & \mathbf{T}_{22} \end{pmatrix} = \begin{pmatrix} \mathbf{S}_{21} - \mathbf{S}_{22}[\mathbf{S}_{12}]^{-1}\mathbf{S}_{11} & \mathbf{S}_{22}[\mathbf{S}_{12}]^{-1} \\ -[\mathbf{S}_{12}]^{-1}\mathbf{S}_{11} & [\mathbf{S}_{12}]^{-1} \end{pmatrix}$$

In certain cases a quasi-analytical approach can be taken (e.g., for a symmetrically truncated beam, useful recursion relationships can be derived).

For a system of cascaded optical components, the transmission matrices are more straightforward to manipulate since  $\mathbf{T}_{total} = \mathbf{T}_N \cdot \dots \cdot \mathbf{T}_i \cdot \dots \cdot \mathbf{T}_2 \cdot \mathbf{T}_1$ .

## Linear Scattering Operators

Unlike a waveguide, a quasi-optical device supports a large (essentially infinite) number of modes at each port. By a careful choice of basis set we can minimize the effective number of modes we need to include.

To include reflections, we can represent a mode traveling in the reverse ( $-z$ ) direction by replacing  $z$  with  $-z$  in the expression for the mode derived earlier. This is equivalent to taking the complex conjugate of the spatial part of the forward-moving mode. For example, in symmetric Laguerre modes (only phase terms are affected),

$$\begin{aligned} E_p^+(r, z, t) &= E_p^+(r, z) \exp(i\omega t) \\ E_p^-(r, z, t) &= [E_p^+(r, z)]^* \exp(i\omega t) \\ E_p^-(r, z) &= [E_p^+(r, z)]^* \end{aligned}$$

In general, we then assume that some **linear scattering** acts on the incident modal fields (such as truncation, aberrational effects, partial reflections, etc.)

$$\begin{aligned} E_{refl} &= L_{11}[E_{inc}] \\ b_{1p} &= \int_S [E_p^-]^* \cdot L_{11}[E_q^+] 2\pi r dr = \langle E_p^- | L_{12}[E_q^+] \rangle \end{aligned}$$

The  $L_{mn}$  are operators representing the transfer function of the system for input at optical port  $n$  and output at optical port  $m$ , acting on the nominated input function

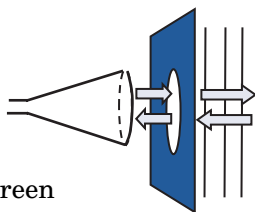
$$\begin{aligned} S_{11}(i, j) &= \langle E_i^- | L_{11}(E_j^+) \rangle & S_{12}(i, j) &= \langle E_i^- | L_{12}(E_j^-) \rangle \\ S_{21}(i, j) &= \langle E_i^+ | L_{21}(E_j^+) \rangle & S_{22}(i, j) &= \langle E_i^+ | L_{22}(E_j^-) \rangle \end{aligned}$$

where  $E^\pm(z)$  are the expansion functions for waves travelling in the  $\pm z$  direction  $E^-(z) = E^+(-z)$ . For example, in symmetric Gaussian–Laguerre modes,

$$\begin{aligned} E_{p0}^\pm(r, z) &= \frac{w_o}{w(z)} L_p \left[ 2 \frac{r^2}{w(z)^2} \right] \exp \left[ -\frac{r^2}{w(z)^2} \right] \\ &\times \exp \left[ \mp i k [z + r^2/2R(z)] \right] \times \exp [\pm (2p + 1) i \phi_0(z)] \end{aligned}$$

## Truncation at an Aperture

Consider truncation of a beam, incident from the left, by a reflecting or absorbing screen with a concentric **aperture** of radius  $r_T$ . Assume no incident field from the right:



$$E = 0 \text{ to the right of screen}$$

$$\Rightarrow E_{refl} = -E_{inc} \text{ over screen}$$

$$E_{trans} = E_{inc} \text{ over aperture}$$

Thus, assuming  $\mathbf{a}_2 = 0$  (see scattering matrix formulism),

$$S_{11}(\mu, \nu) = \left[ - \int_{r > r_T} E_\mu(r, z_0) \cdot E_\nu(r, z_0) \cdot 2\pi r dr \right]$$

Similarly, for the transmitted beam

$$S_{21}(\mu, \nu) = \left[ - \int_{r < r_T} \{E_\mu(r, z_0)\}^* \cdot E_\nu(r, z_0) \cdot 2\pi r dr \right]$$

For the case of an absorbing stop, since there are no reflections, the  $\mathbf{S}_{11}$  and  $\mathbf{S}_{22}$  are identically zero:

$$\mathbf{S} = \begin{bmatrix} \mathbf{S}_{11} & \mathbf{S}_{12} \\ \mathbf{S}_{21} & \mathbf{S}_{22} \end{bmatrix} = \begin{bmatrix} 0 & \mathbf{S}_{12} \\ \mathbf{S}_{21} & 0 \end{bmatrix}$$

$$S_{12}(\mu, \nu) = \int_{ap} [E_\mu^-(x', y')]^* E_\nu^-(x', y') dA$$

$$S_{21}(\mu, \nu) = \int_{ap} [E_\mu^+(x', y')]^* E_\nu^+(x', y') dA$$

where in the case of associated Laguerre modes and a coaxial circular stop,

$$\mathbf{S}_{21}(\mu(\rightarrow pm), \nu(\rightarrow qm)) = S_{qm, pm}(r_T/w) = I_{pqm}(u_T) \times \exp(2i(q-p)\Delta\phi_0),$$

$$\text{where } I_{pqm}(u_T) = \int_0^{u_T} du \left[ \frac{u^m L_p^m(u) L_q^m(u) \exp(-u)}{\sqrt{(p+m)!(q+m)!/p!q!}} \right]$$

$$\text{and } u_T = 2\left(\frac{r_T}{w}\right)^2$$

## Perfect Lenses and Pure Propagation

---

For transmission at **perfect lenses** (or equivalently reflections from curved mirrors), it is convenient to treat the component as part of a beam guide. Then, basis functions at each port correspond to the appropriately transformed functions at the other (using *ABCD* matrices or another simple method). Two sets of basis functions are thus defined in terms of corresponding different beam waist sizes and locations for the input and output (object and image) spaces.

This process defines a set of beam waveguide modes, with respect to which the fields at any plane of constant  $z$  can be expanded.

A lens is not a mode scatterer, and an excitation of a few low-order modes continues to be well approximated with a limited mode set.

On the other hand, for any partial reflections at lens surfaces, this simplification does not apply in general, and many modes are needed. However, in practice, such reflected higher-order modes tend to be truncated.

The reflection coefficient of a **pure propagation path** is zero, and so  $S_{11} = S_{22} = 0$ . There is no scattering between modes, but phase slippages must be taken into account. Consequently,

$$\mathbf{S} = \begin{bmatrix} \mathbf{S}_{11} & \mathbf{S}_{12} \\ \mathbf{S}_{21} & \mathbf{S}_{22} \end{bmatrix} = \begin{bmatrix} 0 & \mathbf{D}_{21} \\ \mathbf{D}_{21} & 0 \end{bmatrix}$$

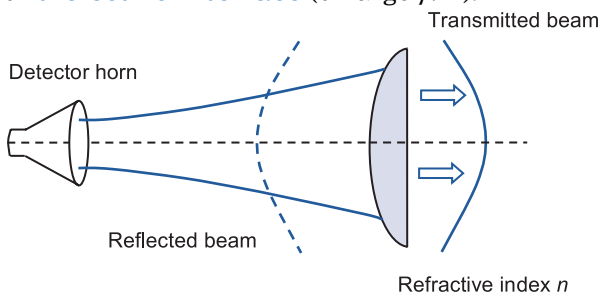
$$\mathbf{D}_{21}(m, n) = \delta_{mn} \exp[i\Delta\phi_m - ik(z_2 - z_1)]$$

$$\mathbf{T} = \begin{bmatrix} \mathbf{T}_{11} & \mathbf{T}_{12} \\ \mathbf{T}_{21} & \mathbf{T}_{22} \end{bmatrix} = \begin{bmatrix} \mathbf{D}_{21} & 0 \\ 0 & [\mathbf{D}_{21}]^{-1} \end{bmatrix} = \begin{bmatrix} \mathbf{D}_{21} & 0 \\ 0 & [\mathbf{D}_{21}]^* \end{bmatrix}$$



## Reflections at Dielectric Interfaces (Lens Surfaces)

Consider a quadratic phase curvature produced by a thin lens or **dielectric interface** (of large  $f/D$ ).



The effective curvature  $1/R_{surf}$  is defined to be positive if the interface appears concave, when looking at it along the direction of the positive  $z$  axis. We can write

$$\mathbf{S} = \begin{bmatrix} \mathbf{S}_{11} & \mathbf{S}_{12} \\ \mathbf{S}_{21} & \mathbf{S}_{22} \end{bmatrix} = \begin{bmatrix} \gamma \mathbf{L}_{11} & (1+\gamma)\mathbf{I} \\ (1+\gamma)\mathbf{I} & \gamma \mathbf{L}_{22} \end{bmatrix}$$

where  $\gamma$  is the reflection coefficient at the interface

$$\gamma = \frac{1-n}{1+n} \quad \text{and}$$

$$\mathbf{L}_{11}(\mu, \nu) = \int_{ap} [E_{\mu}^{-}(r, \phi, z)]^* E_{\nu}^{+}(r, \phi, z) \exp[ikr^2/R_{surf}] dA$$

$$\mathbf{L}_{22}(\mu, \nu) = \int_{ap} [E_{\mu}^{+}(r, \phi, z)]^* E_{\nu}^{-}(r, \phi, z) \exp[-ikr^2/R_{surf}] dA$$

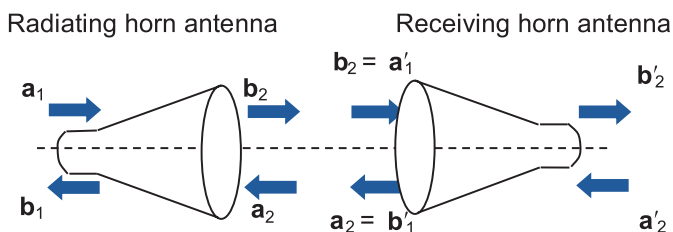
If we are only interested in the coupling to the reflected fundamental mode (the one most likely to couple to a detector), then

$$\begin{aligned} \mathbf{L}_{11}(0,0) &= \int_{ap} \left[ \exp\left(-\frac{r^2}{w^2} + ik \frac{r^2}{2R_{in}}\right) \right]^* \\ &\cdot \left[ \exp\left(-\frac{r^2}{w^2} - ik \frac{r^2}{2R_{in}}\right) \right] \exp[ikr^2/R_{surf}] dA \\ \mathbf{L}_{22}(0,0) &= \int_{ap} \left[ \exp\left(-\frac{r^2}{w^2} - ik \frac{r^2}{2R_{in}}\right) \right]^* \\ &\cdot \left[ \exp\left(-\frac{r^2}{w^2} + ik \frac{r^2}{2R_{in}}\right) \right] \exp[-ikr^2/R_{surf}] dA \end{aligned}$$

## Standing Waves in Horn-Fed Systems

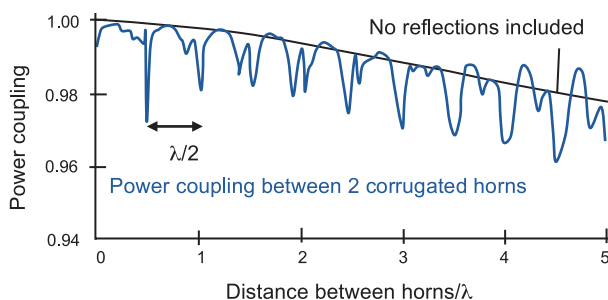
In **horn-antenna-fed systems** we can combine a scattering matrix description of a horn antenna using waveguide modes with a scattering matrix description of a quasi-optical system using Gaussian beam modes, and thus model the **standing waves** between two horn antennas.

We can include in the formulism the scattering matrices that convert antenna [or waveguide (WG)] modes to free-space (FS) modes and vice versa:



$$\mathbf{S}_{21}^{\text{WG} \rightarrow \text{FS}}(\mu, \nu) = \iint_{ap} e_{\mu}^{\text{WG}*} \psi_{\nu}^{\text{FS}} dS$$

In order to include both orthogonal modes, a vector description of propagation in the waveguide is needed.



## Cascading Scattering and Transmission Matrices

We calculate the total **scattering matrix** for a combination of components by **cascading** the individual matrices appropriately, yielding the overall transmission to the output port and reflection back at the input port. Thus, if

$$\mathbf{S}^a = \begin{bmatrix} \mathbf{S}_{11}^a & \mathbf{S}_{12}^a \\ \mathbf{S}_{21}^a & \mathbf{S}_{22}^a \end{bmatrix}; \mathbf{S}^b = \begin{bmatrix} \mathbf{S}_{11}^b & \mathbf{S}_{12}^b \\ \mathbf{S}_{21}^b & \mathbf{S}_{22}^b \end{bmatrix}, \text{ where } \mathbf{S}^b \text{ follows } \mathbf{S}^a, \text{ then}$$

$$\mathbf{S}^c = \begin{bmatrix} \mathbf{S}_{12}^a [\mathbf{I} - \mathbf{S}_{11}^b \mathbf{S}_{22}^a]^{-1} \mathbf{S}_{11}^b \mathbf{S}_{21}^a + \mathbf{S}_{11}^a & \mathbf{S}_{12}^a [\mathbf{I} - \mathbf{S}_{11}^b \mathbf{S}_{22}^a]^{-1} \mathbf{S}_{12}^b \\ \mathbf{S}_{21}^b [\mathbf{I} - \mathbf{S}_{22}^a \mathbf{S}_{11}^b]^{-1} \mathbf{S}_{21}^a & \mathbf{S}_{21}^b [\mathbf{I} - \mathbf{S}_{22}^a \mathbf{S}_{11}^b]^{-1} \mathbf{S}_{22}^a \mathbf{S}_{12}^b + \mathbf{S}_{22}^b \end{bmatrix}$$

By contrast, **transmission matrices** are extremely useful for a cascading system, as we simply multiply the **T** matrices together (which we cannot do for the **S** matrices). Note that  $[\mathbf{S}_{12}]^{-1}$  can be singular in **T**. Nevertheless, both formulisms (whether cascading **S** or **T**) are identical, and

$$\mathbf{T}^c = \begin{bmatrix} \mathbf{T}_{11}^b & \mathbf{T}_{12}^b \\ \mathbf{T}_{21}^b & \mathbf{T}_{22}^b \end{bmatrix} \cdot \begin{bmatrix} \mathbf{T}_{11}^a & \mathbf{T}_{12}^a \\ \mathbf{T}_{21}^a & \mathbf{T}_{22}^a \end{bmatrix}$$

$$= \begin{bmatrix} \mathbf{T}_{11}^b \mathbf{T}_{11}^a + \mathbf{T}_{12}^b \mathbf{T}_{21}^a & \mathbf{T}_{11}^b \mathbf{T}_{12}^a + \mathbf{T}_{12}^b \mathbf{T}_{22}^a \\ \mathbf{T}_{21}^b \mathbf{T}_{11}^a + \mathbf{T}_{22}^b \mathbf{T}_{21}^a & \mathbf{T}_{21}^b \mathbf{T}_{12}^a + \mathbf{T}_{22}^b \mathbf{T}_{22}^a \end{bmatrix}$$

For example, if we consider a series of absorbing stops with no reflections, we can calculate the overall scattering matrix for the system by cascading sequential truncation and propagation matrices:

$$\mathbf{T}^{total} = \mathbf{T}^N \dots \mathbf{T}^c \cdot \mathbf{T}^b \cdot \mathbf{T}^a = \mathbf{T}^N \dots \mathbf{T}_{stop}^c \cdot \mathbf{T}_{prop}^b \cdot \mathbf{T}_{stop}^a$$

$$= \begin{bmatrix} \mathbf{A}_{21}^N & 0 \\ 0 & [\mathbf{A}_{12}^N]^{-1} \end{bmatrix} \dots \begin{bmatrix} \mathbf{A}_{21}^c & 0 \\ 0 & [\mathbf{A}_{12}^c]^{-1} \end{bmatrix} \cdot \begin{bmatrix} \mathbf{D}_{21}^b & 0 \\ 0 & [\mathbf{D}_{21}^b]^* \end{bmatrix} \cdot \begin{bmatrix} \mathbf{A}_{21}^a & 0 \\ 0 & [\mathbf{A}_{12}^a]^{-1} \end{bmatrix}$$

$$= \begin{bmatrix} \mathbf{A}_{21}^N \dots \mathbf{A}_{21}^c \mathbf{D}_{21}^b \mathbf{A}_{21}^a & 0 \\ 0 & [\mathbf{A}_{12}^N]^{-1} \dots [\mathbf{A}_{12}^c]^{-1} [\mathbf{D}_{21}^b]^* [\mathbf{A}_{12}^a]^{-1} \end{bmatrix}$$

Thus, for the system as a whole,

$$\mathbf{T}_{11}^{total} = \mathbf{A}_{21}^N \dots \mathbf{A}_{21}^c \mathbf{D}_{21}^b \mathbf{A}_{21}^a$$

$$\mathbf{T}_{22}^{total} = [\mathbf{A}_{12}^N]^{-1} \dots [\mathbf{A}_{12}^c]^{-1} [\mathbf{D}_{21}^b]^* [\mathbf{A}_{12}^a]^{-1}$$

## Off-Axis Mirrors (Ellipsoidal)

Refocusing (curved) mirrors are commonly employed for recollimation in quasi-optical beam guides. They are also useful for feeding a divergent beam from a source into a beam guide or focusing a beam onto a detector. Normally, mirrors are used in an **off-axis configuration** to prevent the beam from interfering with itself. To minimize beam distortion, the angle of incidence  $i$  of the beam should be less than 45 deg (bending angle  $2i < 90$  deg). However, in telescopes and large antenna reflectors, an on-axis Cassegrain configuration is often encountered.

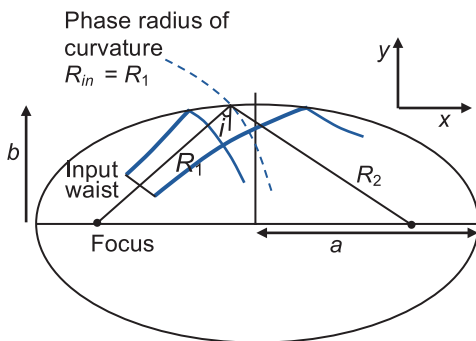
A reflecting surface in the form of an **ellipsoid of revolution** around its major ( $x$ ) axis correctly transforms the phase radius of curvature of an incident Gaussian beam  $R_{in} = R_1$  into a phase radius of curvature for the reflected beam of  $R_{out} = -fR_{in} / (R_{in} - f) = -R_2$ , where  $f$  is the focal length of the mirror. The major and minor axes and eccentricity of the ellipsoid are given by

$$a = (R_1 + R_2)/2, \quad b = \sqrt{R_1 R_2} \cos i \text{ and}$$

$$e = \sqrt{1 - b^2/a^2} = \sqrt{1 - 4R_1 R_2 \cos^2 i / (R_1 + R_2)^2}$$

where  $x^2/a^2 + y^2/b^2 = 1$ , and

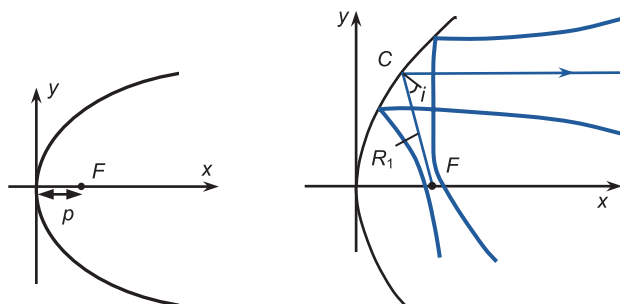
$$f = R_1 R_2 / (R_1 + R_2)$$



Note that the waists do not coincide with the geometrical foci of the ellipsoid.

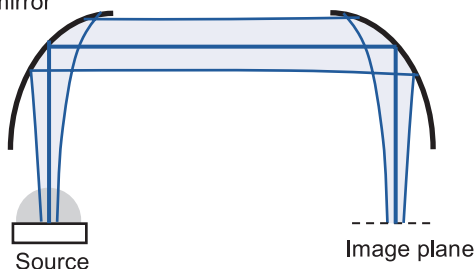
## Off-Axis Mirrors (Parabolic)

If either the incident or the reflected beam is required to have a waist at the mirror, then the correct surface is a **paraboloid of revolution** around the axis of symmetry of a (parent) parabola ( $x$  axis in the figure below), and the focal length of the mirror must match the phase front radius of curvature  $f = R_1$  of the noncollimated beam. If the equation for the parabola is given by  $y^2 + z^2 = 4px$ , where  $p$  is the distance from the vertex to the focus  $F$  of the parabola, then  $p = R_1 \cos^2 i$ , where  $i$  is the angle of incidence at the mirror. The intersection point of the axis of the reflected beam with the mirror surface is given by  $x_C = R_1 \sin^2 i$ , and  $y_C = R_1 \sin 2i$ .



If either the input or output waist in a focusing mirror setup is wide ( $\gg \lambda$ ), causing the corresponding phase radius of curvature to be very large at the mirror (i.e.,  $R \gg f$ ), for convenience, a parabolic reflector can be used to replace an ellipsoid in quasi-optical systems, particularly for **Gaussian beam telescope** (GBT) arrangements.

Parabolic mirror

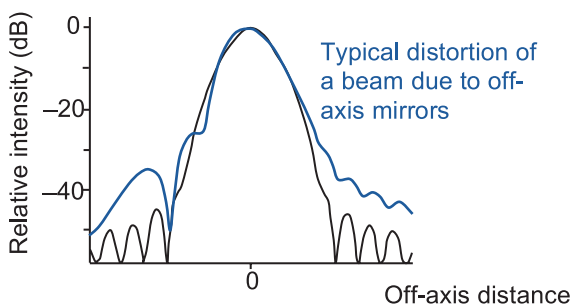


## Off-Axis Mirrors (Distortion and Cross-Polarization)

Because of projection effects, the reflected beam from an off-axis mirror suffers from **amplitude distortion (beam squint)**. For a reflected Gaussian beam, the maximum is no longer on the propagation axis; however, the Gaussicity of such a beam is still high for long focal ratios and small angles of incidence  $i$ . The Gaussicity (the fractional power coupling to a pure Gaussian, i.e., the fundamental mode) is given by

$$K_G = 1 - w_m^2 \tan^2 i / (8f^2)$$

where  $w_m$  is the beam waist at the mirror, which has a focal length  $f$ . In a two-mirror system, the distortion can be compensated for with the correct arrangement, depending on whether the intervening waist is narrow (i.e., a focused beam) or wide (i.e., a collimated beam). If the mirror does not have the correct ellipsoid of revolution shape, then phase distortion also occurs at the mirror, leading to additional phase aberration effects.



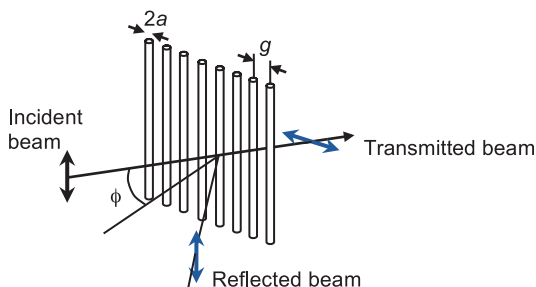
The fraction of power in an incident linearly polarized beam that is scattered into the **cross-polar** direction is given by

$$K_{Xpol} = [w_m \tan i / (2f)]^2$$

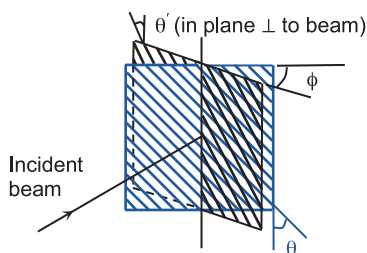
The cross-polar field has the approximate form of an asymmetric higher-order Hermite–Gaussian mode with two lobes along the direction of the incident linear polarization.

## Polarizing Grids

**Polarizing grids** consisting of parallel conducting wires are used to linearly polarize a beam. If the wire spacing  $g < \lambda/4$ , and the wire diameter  $2a < \lambda/10$ , the grid effectively transmits an electric field that is polarized perpendicular to the wires and reflects a field polarized parallel to the wires, achieving good polarization discrimination. A wire grid inclined at an angle  $\phi$  can be used to combine beams or separate one beam into two beams with orthogonal polarization.

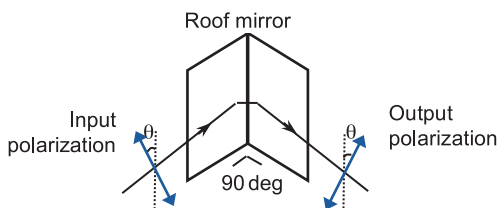


The grid wires themselves can also be tilted at an angle  $\theta$ , leading to projection effects. The polarization directions of the transmitted and reflected fields are given by  $\beta_{trans} = \pi/2 - \theta'$  and  $\beta_{refl} = \theta'$ , with respect to the vertical (see below). For an incident field itself polarized at angle  $\alpha$  to this vertical direction, the fractions of the power transmitted and reflected are given by  $T_{trans} = \sin^2(\alpha - \theta')$  and  $T_{refl} = \cos^2(\alpha - \theta')$ , respectively. If the grid is also rotated through an angle  $\phi$ ,  $T_{trans}$  and the polarization directions of the transmitted and reflected beams are complex functions of  $\alpha$ ,  $\theta$ , and  $\phi$ .



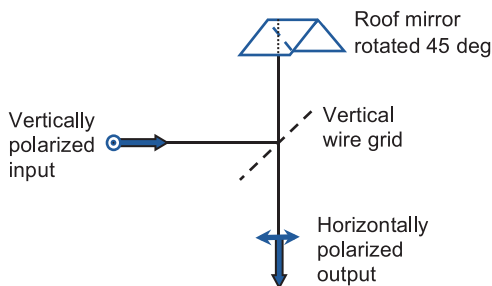
## Roof Mirrors as Polarization Rotators

A **roof mirror** (corner reflector), which consists of two flat reflecting surfaces at right angles to each other, reflects a beam back along the incident direction while at the same time rotating the angle of polarization of a linearly polarized beam, as viewed along the axis of propagation toward the mirror. If we define the axis of the roof mirror as parallel to the intersection line between the two surfaces, then an incident beam linearly polarized at an angle  $+\theta$  is reflected with its direction of polarization rotated to  $-\theta$ .



This is in contrast to normal reflection from a single plane-reflecting surface in which the direction of polarization is unchanged.

If one chooses  $\theta = 45$  deg, then the field is rotated by 90 deg with respect to the original direction, a useful property when combined with polarizing grids. For example, a roof mirror–grid combination can rotate the polarization angle of a beam by 90 deg.



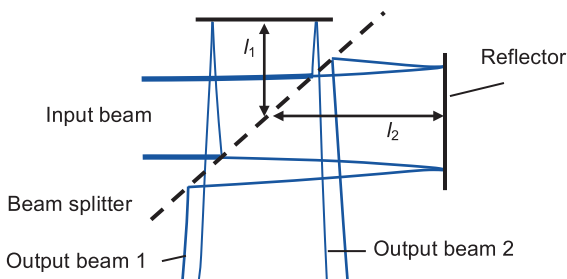
Another critical application of roof mirror–grid combinations is in **path length modulators**, especially in polarizing **Michelson interferometers**.



## Dual-Beam Interferometers (Tunable Filters)

**Dual-beam interferometers**, which work on either an amplitude or polarization division scheme, are useful as **tunable filters**.

In particular, the Michelson interferometer is a very convenient configuration for application in quasi-optical systems. In the amplitude splitting configuration, the path length difference is  $\Delta = 2(l_2 - l_1)$ , where  $l_1$  and  $l_2$  are the path lengths to the two reflectors from the beam splitter. The fractional power transmission of a monochromatic signal of wavelength  $\lambda$  from input to output is:  $|T|^2 = 2|r|^2|t|^2[1 + \cos(k\Delta)]$ , where  $k = 2\pi/\lambda$ , and  $r$  and  $t$  are the complex reflection and transmission amplitude coefficients, respectively. Transmission is a maximum when  $\Delta_{max} = M\lambda$  and a minimum when  $\Delta_{min} = (M + 1/2)\lambda$ , where  $M$  is an integer. For  $|T|^2 = 1$ , we require that  $|r|^2 = |t|^2 = 1/2$ .



For long-wavelength beams, diffraction must be taken into account, and a mismatch occurs between the recombined beams both in the beamwidths and the phase slippages (because of the different distances traveled).

A Michelson interferometer can be used as a **sideband filter** in a heterodyne system by having the transmission maximum for one sideband coincide with the transmission minimum for the other. This coincidence occurs when the path length difference is  $\Delta_{SSB} = (M - 1/2)(\lambda_{IF}/2)$ .

## Diffraction Losses in Interferometers

The issue of **diffraction losses** arises because the beam in an interferometer is not a plane wave. Thus, effectively, two Gaussian beams that have travelled different path lengths are combined at the output, each having suffered a different amount of diffraction. Ideally, the path length difference should be less than the confocal distance; otherwise, the interferometer will not behave as expected.

In order to minimize these effects, the beam should not have spread out too much. In this case the mismatch in amplitude and phase radius of curvature is not large, and the phase-slippage terms for the two paths are not significantly different.

In terms of the recombined beams with complex curvature terms  $q_1$  and  $q_2$ ,

$$E_{out} = E_0 \left[ \frac{\exp(-ikr^2/2q_1)}{q_1} + \frac{\exp(-ikr^2/2q_2)}{q_2} \right]$$

If an output device couples perfectly to one of the Gaussian beams, then the power coupling coefficient to the total output (sum of the two fields) from the interferometer is

$$K = |rt|^2 \left[ \frac{2 + \alpha^2 + 2[\cos k\Delta + \alpha \sin k\Delta]}{1 + \alpha^2} \right]$$

where the **diffraction parameter**  $\alpha$  is defined as

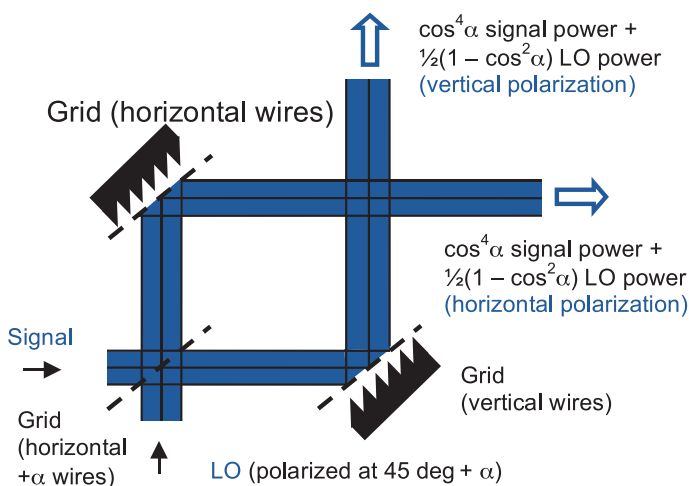
$$\alpha = \frac{\lambda\Delta}{2\pi w_0^2} = \frac{\Delta}{2z_c}$$

$z_c$  is the confocal distance, and  $\Delta$  is the path length difference through the interferometer.

Clearly, in order for the power coupling to be close to ideal, we require that  $\Delta \ll 2z_c$ . This sets the limit on the path for a given beam diameter.

## Diplexers and Multiplexers

Two or more beams can be coupled together using quasi-optical techniques. A typical application is coupling to a local oscillator signal in a heterodyne (radio-like) receiver for high-resolution spectroscopy. The coupling of two beams can be achieved using a simple beam splitter that consists of a partially reflecting/transmitting surface or a polarizing grid arrangement. However, such systems are lossy.



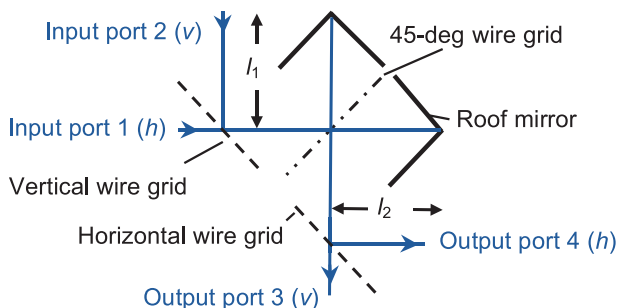
Diplexer based on polarization splitting [courtesy of ESA (HIFI)].

If the two beams are separated in frequency, relatively lossless coupling can be achieved using interferometric techniques (in particular, **Fabry–Pérot**, Michelson, and **Martin–Puplett** interferometers have been applied for this purpose).

In general, for multiplexing several beams (array systems), phase gratings can be designed, while dichroics are also utilized for splitting an incoming signal into several wavelength bands.

### Four-Port Dual-Beam Interferometer (Diplexer)

Both amplitude and polarization versions of **four-port dual-beam interferometers** exist. These interferometers are extremely valuable as signal diplexers with two input ports for the two signals to be combined (two output ports are also possible). In the polarizing Michelson interferometer, horizontally polarized radiation from input port 1 is combined with vertical radiation from port 2 using a vertical wire grid. The two beams are then incident on the 45-deg (in projection) wire grid, which splits them. The appropriately orientated roof mirrors rotate the angles of polarization so that the beams recombine at the 45-deg grid and are incident on a horizontal wire grid, which splits the output beam into two polarization components (ports 3 and 4).



The path length difference is  $\Delta = 2(l_1 - l_2)$ , and the fractional power transmission to the two output ports is

$$|T|^2 = [1 + \cos(k\Delta)]/2 \text{ and } |T|^2 = [1 - \cos(k\Delta)]/2$$

For a diplexer in a heterodyne system, if an LO is fed through port 2, for example, and the signal through port 1, then both exit through the same port (and are co-polarized) if

$$\Delta = (M - 1/2)(\lambda_s - \lambda_{LO}) = (M - 1/2)(\lambda_{IF})$$

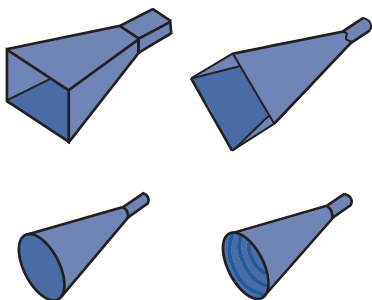
where  $\lambda_{IF}$  is the wavelength associated with the intermediate frequency, and  $M$  is an integer. The frequency of successive transmission bands is given by

$$\delta\nu = c/\Delta = \nu_{IF}/(M - 1/2)$$

## Horn Antenna Feeds

**Horn antennas** are used as feed systems for launching and detecting THz waves, particularly at longer wavelengths. They produce beams that have low sidelobe levels and are well approximated by a Gaussian distribution. A horn antenna can be regarded as a flared section of waveguide that acts as a smooth transition to free space.

The far-field beamwidth of the horn is dependent on the size of the aperture and the phase error across the horn flare ( $s = d^2/8\lambda L$ , where  $d$  is the diameter or side length of the horn aperture, and  $L$  is the slant length of the horn). If the flare is gentle, then the field structure at the horn aperture is readily predicted from waveguide theory.



In terms of feeding quasi-optical systems, the most important parameters are the effective waist size and the location of the Gaussian beam that best fits the radiation field. To a good approximation, the field at the aperture has a well-defined spherical phase front, with a radius of curvature  $R$ . For the typical linearly flared conical and pyramidal shapes encountered, the center of curvature of the phase front is located at the apex of the horn, i.e.,  $R = L$ . The phase variation can be expressed using the usual quadratic approximation for paraxial beams  $\exp(-ikr^2/2L)$ .

The **sidelobe** structure is determined by the high-frequency structure of the horn aperture fields (such as sharp truncation of the field or its derivatives at the edge of the aperture).

## Corrugated Conical Horns (Scalar Feed)

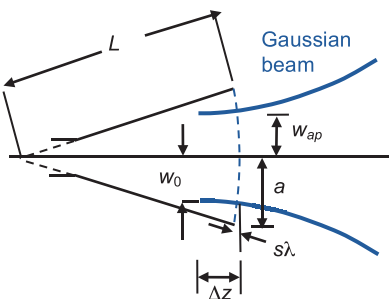
A **corrugated conical horn** (or “scalar feed”), which operates under the balanced hybrid condition with corrugations a quarter-wavelength deep at the aperture, produces a hybrid  $HE_{11}$  mode of the form:

$$\mathbf{E}_{ap} = J_0 \left( 2.405 \frac{r}{a} \right) \exp \left( -ik \frac{r^2}{2L} \right) \mathbf{j}$$

where  $a$  is the radius of the horn aperture,  $L$  is its slant length, and the field is polarized in the  $\mathbf{j}$  direction. For the best-fit Gaussian at the horn aperture, the beamwidth parameter is given by  $w_{ap} = 0.6435a$ , so that the waist size and its position behind the horn aperture are:

$$w_0 = \frac{0.6435 a}{\sqrt{1 + \left( \frac{\pi (0.6435 a)^2}{\lambda L} \right)^2}}$$

$$\Delta z = \frac{L}{1 + \left( \frac{\lambda L}{\pi (0.6435 a)^2} \right)^2}$$



The higher-mode content (amplitude and power) in terms of symmetric Laguerre–Gaussian beam modes  $L_p^0$  is given in the table.

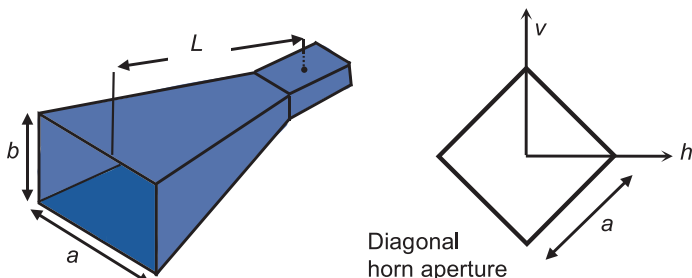
Mode No.	Amplitude	Power
0	0.9903	0.9807
1	0	0.0000
2	-0.1204	0.0145
3	-0.0432	0.0019
4	0.0194	0.0004
5	0.0340	0.0012
6	0.0199	0.0004
7	0.0002	0.0000
8	-0.0125	0.0002
9	-0.0152	0.0002

## Smooth-Walled Horns (Pyramidal and Diagonal)

A **pyramidal horn** antenna (of aperture width  $a$  and height  $b$ ) produces a rectangular waveguide TE<sub>10</sub> fundamental mode field at its aperture with a phase error given by the slant lengths ( $L_x$  and  $L_y$ ) of the horn:

$$\mathbf{E}_{ap} = E_0 \cos\left(\frac{\pi x}{a}\right) \exp\left(-ik \left[ \frac{x^2}{2L_x} + \frac{y^2}{2L_y} \right]\right) \hat{\mathbf{j}}, \quad |x| \leq a/2, |y| \leq b/2$$

For optimum operation as a quasi-Gaussian feed, we require that  $L_x = L_y$ , and  $b/a = 0.7$ , giving a symmetric best-fit Gaussian with a width parameter  $w_{ap} = 0.35a = 0.5b$  at the horn aperture.



A **diagonal horn** is an improvement on a rectangular **smooth-walled horn** (pure TE<sub>10</sub>) that produces a beam pattern with the same E- and H-plane patterns. It is relatively easy and thus inexpensive to manufacture (no corrugations). A flared square horn section is rotated by 45 deg with respect to a rectangular waveguide feeding it with a single TE<sub>10</sub> mode. Two orthogonal TE<sub>10</sub> modes are thus excited in the horn with equal phase and amplitude. For a horn of side length  $a$  and slant length  $L$ , the electric field at the aperture is given by

$$\mathbf{E}_{ap} = \left[ \hat{\mathbf{i}} \cos\left(\frac{\pi y}{a}\right) + \hat{\mathbf{j}} \cos\left(\frac{\pi x}{a}\right) \right] \exp\left(-ik \left[ \frac{x^2 + y^2}{2L} \right]\right) \quad |x| \leq \frac{a}{2}; |y| \leq \frac{a}{2}$$

This yields a quasi-Gaussian field with beamwidth parameter  $w_{ap} = 0.43a$  at the aperture.

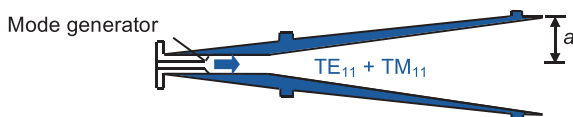
## Conical Smooth-Walled Horns (Single and Dual Mode)

A **smooth-walled conical horn** antenna (of radius  $a$  and slant length  $L$ ) produces a cylindrical waveguide  $TE_{11}$  fundamental mode field of the form:

$$\mathbf{E}_{ap} \propto \{[J_0(\chi r/a) + J_2(\chi r/a) \cos 2\phi] \mathbf{i} + [J_2(\chi r/a) \sin 2\phi] \mathbf{j}\} \exp(-ikr^2/2L), \quad \chi = 1.841$$

For operation as a quasi-Gaussian feed, the best-fit Gaussian has beam radius parameter  $w_{ap} = 0.76a$  at the horn aperture.

A **dual-mode horn** is an improvement on a smooth-walled conical horn (pure  $TE_{11}$ ). It produces a nearly symmetric tapered field in a circular horn by also exciting the  $TM_{11}$  mode (with correct amplitude and phase relative to  $TE_{11}$ ) with a carefully designed mode launcher (step plus phasing section) in the throat of the horn.



$$\mathbf{E}_{ap} = \frac{\mathbf{E}(TE_{11}) + \beta \mathbf{E}(TM_{11})}{1 + \beta}$$

The radial component of the field vanishes at the aperture boundary if  $\beta = 0.785$  and the two modes are in phase at the horn aperture. Such horns are relatively easy and inexpensive to manufacture compared to corrugated horns. The input waveguide only allows  $TE_{11}$  propagation; it is the step to a wider waveguide that generates the  $TM_{11}$  mode (the wider guide is above the cutoff for  $TM_{11}$ ). The phasing section brings the two modes into correct phase at the horn aperture. The resulting phase shift imposes limitations on the operational bandwidth.

The fraction of total energy in the co-polar field is 98.6%, with the beamwidth parameter of the best-fit Gaussian beam at the horn aperture given by  $w_{ap} = 0.59a$ .

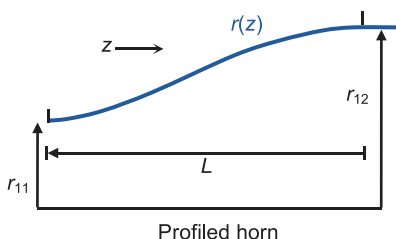


## Shaped Horns and Multimode Feeds

**Shaped horn** designs have been developed for two contrasting requirements. Important examples include (a) profiled corrugated horns that have a reduced length compared to linearly tapered horns, while still providing high gain (although somewhat degraded sidelobe levels), and (b) so-called super-Gaussian corrugated horns that are shaped in the throat region in order to generate higher-order hybrid waveguide modes with a long parallel section for phasing the modes, so that a more perfect Gaussian taper is produced at the horn aperture (Gaussicity of 99.5%), and (c) flared horns that also produce a higher Gaussicity by shaping the horn close to the aperture.

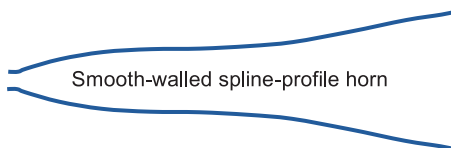
A suitable form for the profile corrugated horn is given by

$$r(z) = r_{11} + (r_{12} - r_{11}) \times \left[ \frac{z}{L} (1 - A) + A \sin^2 \left( \frac{\pi z}{2L} \right) \right]$$



Parameter  $A$  determines the amount of profiling.

A value of  $A = 0.7$  gives a nonlinear yet gradual change in diameter to ensure adequate performance. Smooth-walled spline-profile horns have also been developed as an alternative to the corrugated horn for wideband millimeter-wave applications:

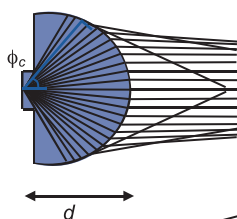


For horns fed by an **over-moded waveguide** feeding an incoherent detector (e.g., a bolometer), many higher-order waveguide modes can also propagate. In this situation determining the horn beam pattern is quite complex because the beam is only **partially coherent**.

## Lens Antennas

**Lens antennas** tend to be used at higher frequencies in the THz band when the manufacture of horn antennas becomes problematic.

A planar antenna typically acts as the basic radiating source that produces a highly divergent beam. The lens then quasi-collimates this beam to produce a far-less-divergent far-field pattern, which is useful for input to quasi-optical systems. The lens material usually matches the substrate material used for the planar antenna (typically high-resistivity silicon). The figure shows ray-tracing designs of a collimating lens (top) and a hyperhemispherical lens (bottom).



The distance between the focal point and the apex of a collimating lens is given by

$$d_{coll} = R \left( \frac{n}{n-1} \right)$$

and for a hyperhemispherical lens by

$$d_{hyper} = R \left( \frac{n+1}{n} \right)$$

where  $n$  is the refractive index of the lens, and  $R$  is the radius.

A collimating lens suffers loss due to the total internal reflections for rays from the source approaching the critical angle for total internal reflection  $\theta_c = \sin^{-1}(1/n)$ , which defines an effective aperture for the lens. At larger angles, rays are trapped within the lens. The critical emission angle is given by

$$\phi_c = \sin^{-1} \left( \frac{n-1}{n} \right)$$

## System Design

In many cases, we wish to design a quasi-optical system that couples one beam to another. If the desired input and output waists are  $w_{0,in}$  and  $w_{0,out}$ , then the **magnification** of the system is given by  $M = w_{0,out}/w_{0,in}$ . If the separation of the waists is a free parameter, then

$$d_{in} = f \pm M^{-1} (f^2 - f_0^2)^{0.5} \quad \text{and} \quad d_{out} = f \pm M (f^2 - f_0^2)^{0.5}$$

where

$$f_0 = \frac{\pi w_{0,in} w_{0,out}}{\lambda}$$

We are free to choose the focal length of the system, but it must have a minimum value of  $f = f_0$ . The choice of  $f$  determines the input and output distances, and in the case where  $f = f_0$ ,  $d_{in} = d_{out} = f$  (independent of wavelength). Often, however, the distance between the waists,

$$d = d_{in} + d_{out},$$

is fixed, and therefore so is the focal length. For  $M \neq 1$ ,

$$f = \pm \frac{\left[ (M - M^{-1})^2 f_0^2 + d^2 \right]^{0.5} (M - M^{-1}) - 2d}{(M - M^{-1})^2}$$

while for  $M = 1$ ,

$$f = \frac{d}{4} + \frac{f_0^2}{d}$$

Once the component focal lengths and separations have been fixed, a simple *ABCD* matrix analysis of an equivalent on-axis system (approximating the focusing elements by thin lenses of focal length  $f$ ) can be performed in order to determine the beam waist  $w$  and radius of curvature  $R_{in}$  and  $R_{out}$  at each optical component. The required component sizes depend on the level of beam truncation that can be tolerated, but component diameters of  $4w$  (giving  $-35$  db edge taper) are typical. If off-axis geometry is required, then the angle of throw needed at each component and the radii of curvature can be used to define the reflecting surface.

## Modeling Techniques

---

Optical modeling is concerned with calculating the electromagnetic field over a surface when the field—or currents—over some other surface is known. Techniques such as the **method of moments** attempt to precisely calculate the current distribution over a surface, but in practice approximations often have to be made. In the physical optics approximation, when a field is incident upon an aperture, it is assumed that the field over the opaque region is zero, and the field over the transparent region is the same as the case where there is no aperture. This is reasonable when the radius of curvature of a reflector is many wavelengths, but is not valid at an edge. The **geometrical theory of diffraction** is often additionally used to estimate the effect of aperture edges.

Propagation of the field onto the next optical component (solving the wave equation) requires **diffraction integrals** to be calculated for each field point. The electromagnetic field over an input surface is a vector field, but, when considering narrow-angle paraxial beams, it is often sufficient to consider the components of the vector field separately. Rather than evaluating diffraction integrals directly, it is possible to decompose an assumed source field into modes, each a solution of the wave equation. Propagation to the next optical surface simply involves recombining scaled modes with an appropriate mode-dependent phase-slippage term included. Commonly used mode sets include Gaussian beam modes, **Gabor modes**, and **plane waves**. A plane-wave analysis has the significant advantage that it does not need to be limited to paraxial fields.

At optical wavelengths, away from boundary shadows and abrupt changes in intensity, energy can be considered to be transported along curves, or rays, that obey geometrical laws. This **ray tracing** (geometrical optics) can also be accurate for systems that are highly over-moded. In the THz regime, however,  $\lambda$  is typically an appreciable fraction of component size, and systems tend to be at most few-moded, so diffraction cannot be neglected.

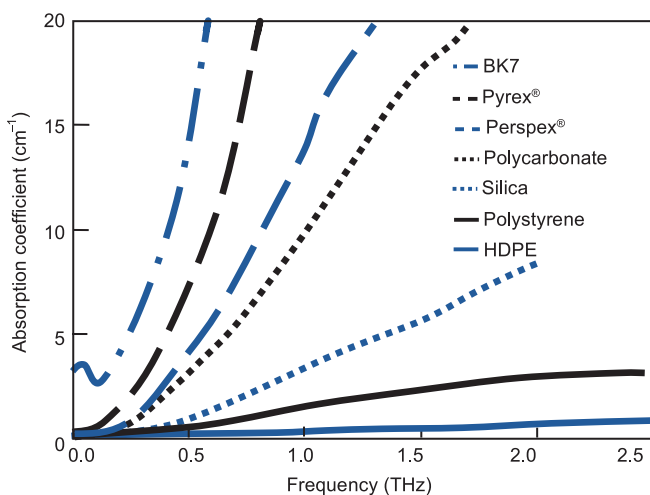
## THz Imaging

THz imaging is a unique tool with many interesting applications. Both **active** and **passive** imaging systems have been developed.

Passive imagers detect the THz radiation emitted or scattered (reflected) by an object, making the object visible through contrast with its surroundings.

In active imaging, a THz source (usually coherent) illuminates the object being imaged, and the scattered radiation from the source is detected. Pulsed-imaging systems based on time-domain THz generation and detection have also been developed, and reflections from buried surfaces have been found and imaged.

THz imaging is particularly useful because in general, dry insulating materials are transparent at these wavelengths.



An example of material absorption measurements taken from Naftaly and Miles (2007).

### THz Imaging (cont.)

---

A significant advantage over X rays is the non-ionizing nature of T rays. Additionally, it is easy to detect dielectric materials due to edge diffraction and partial-absorption effects in T rays. Thus, it is possible to inspect sealed packages for contrasting metal and plastic objects, etc. On the other hand, THz imaging uses more-compact systems than millimeter-wave imaging while also providing higher resolution.

Industrial applications are in nondestructive testing and process control; for example, testing pharmaceutical tablets for integrity, etc. In biomedical applications, THz imaging has been used to detect skin cancers and is proposed for inspection of wound healing; however, one limitation here is that the resolution required is only possible with near-field techniques.

Many systems form images using scanning techniques, which can be time consuming. The development of room-temperature detector arrays will be necessary for real-time THz imaging to reach its full potential.

## THz Spectroscopy

---

Many materials have unique spectral fingerprints in the THz range. Rotational transitions of gaseous molecules result in narrow absorption lines that can be used to identify the molecular structure and the presence of particular species in, for example, the Earth's atmosphere and the interstellar medium.

The THz region of the spectrum is also of interest for detecting and identifying explosives. Here, absorption due to the motion of molecules as a whole can be used as a characterization tool. Many different types of explosives have unique spectral signatures in the THz region. Their strong absorption bands differentiate them from each other and from any concealing material. The noninvasive and nonionizing nature of THz spectroscopy is vital from the safety point of view.

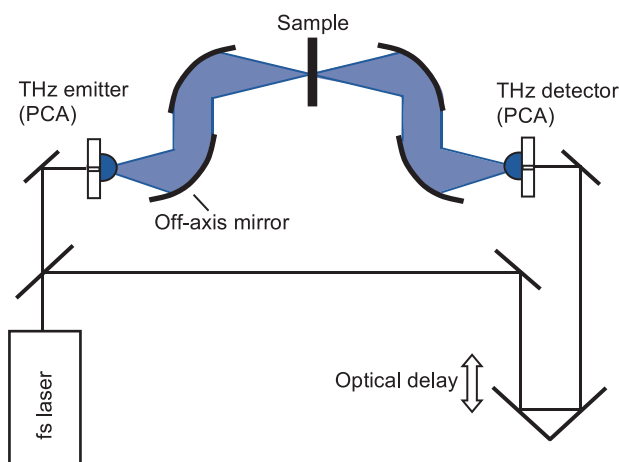
THz spectroscopy is sensitive to both inter- and intramolecular vibrations in different chemical species (e.g., DNA) and can be used to investigate the crystalline state of drugs (polymorphism). Proteins can rearrange their molecular structure on picosecond timescales that can be probed by THz radiation.

In material science, THz spectroscopy can be used to study carrier dynamics in semiconductors and superconductors.

Since many optically opaque materials are transparent to THz radiation, items of interest can be observed through packaging or clothing.

## THz Time-Domain Spectroscopy

**THz time-domain spectroscopy (THz-TDS)** is a spectroscopic technique in which the properties of a material are probed with short pulses of THz radiation. Typically, a pulsed femtosecond laser is used to generate THz pulses of a few picoseconds in a biased PCA. The pulse contains frequency components covering the THz band. For detection, the electrical field of the THz pulse interacts with the same short laser pulse in the PCA detector and produces an electrical signal that is proportional to the THz electric field. Substrate lens antennas are used to collimate the source and detector beams.



By repeating this procedure and varying the timing of the gating laser pulse, it is possible to scan the THz pulse and construct its electric field as a function of time. Subsequently, a Fourier transform is used to extract the frequency spectrum from the time-domain data. By raster scanning the target, it is possible to achieve pulsed terahertz imaging. By its nature, pulsed terahertz imaging gives rise to flexible modes of visualization.



## Equation Summary

---

### Beer–Lambert Law:

$$\tau = \frac{I}{I_0} = \exp(-\alpha x)$$

### Absorbance:

$$A = -\log_{10} \left( \frac{I}{I_0} \right) = \frac{\alpha x}{\ln(10)}$$

### Blackbody radiation:

$$I(\nu, T) = \frac{2h\nu^3}{c^2} \frac{1}{e^{h\nu/k_B T} - 1}$$

### Reflected and transmitted field amplitude:

s-polarization

$$\frac{E_{s,r}}{E_{s,i}} = \frac{n_1 \cos \theta_1 - n_2 \cos \theta_2}{n_1 \cos \theta_1 + n_2 \cos \theta_2} \quad \frac{E_{s,t}}{E_{s,i}} = \frac{2n_1 \cos \theta_1}{n_1 \cos \theta_1 + n_2 \cos \theta_2}$$

p-polarization

$$\frac{E_{p,r}}{E_{p,i}} = \frac{n_2 \cos \theta_1 - n_1 \cos \theta_2}{n_2 \cos \theta_1 + n_1 \cos \theta_2} \quad \frac{E_{p,t}}{E_{p,i}} = \frac{2n_1 \cos \theta_1}{n_2 \cos \theta_1 + n_1 \cos \theta_2}$$

### Snell's law:

$$n_1 \sin \theta_1 = n_2 \sin \theta_2$$

### Reflectivity and transmittance:

$$R = \frac{|E_r|^2}{|E_i|^2} \quad \text{and} \quad T = \frac{n_2 \cos(\theta_2) |E_t|^2}{n_1 \cos(\theta_1) |E_i|^2}$$

### Resonant frequency:

$$\nu = 1/2\pi\sqrt{1/LC}$$

## Equation Summary

---

### Optical rectification, walk-off length:

$$l_w = c\tau_p / (n_{THz} - n_{optical})$$

### Optical rectification, coherence length:

$$l_c = \frac{c}{2\nu_{THz} |n_{optical} - n_{THz}|}$$

### Cyclotron resonance frequency:

$$\omega = \frac{eB}{\gamma m_e}$$

### Synchrotron radiation cone angle:

$$\Delta\theta = \frac{1}{\gamma}$$

### Wavelength of FEL radiation:

$$\lambda_{FEL} = \frac{\lambda_w}{2\gamma^2} (1 + K^2), \quad K = \frac{eB\lambda_w}{2\pi m_e c^2}$$

### Angular dispersion of Smith–Purcell radiation:

$$\lambda = \frac{l}{|n|} \left( \frac{1}{\beta} - \sin\theta \right)$$

### Noise equivalent power:

$$NEP = \frac{n}{R_V \sqrt{\Delta\nu}}$$

$$NEP = \frac{\text{Noise Voltage}/\sqrt{Hz}}{R_V}$$

### Signal-to-noise ratio:

$$SNR = \frac{P_{signal}}{P_{noise}}$$

## Equation Summary

---

$$SNR_{dB} = 10 \log_{10} \frac{P_{signal}}{P_{noise}}$$

$$SNR = \frac{n_{signal}}{n_{noise}}$$

$$SNR = \frac{P_{signal}}{NEP \sqrt{\Delta\nu}}$$

in the case of shot noise:

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

**Specific detectivity:**

$$D^* = \frac{A^{1/2}}{NEP}$$

**Thermal noise generated in a resistor:**

$$V_{rms} = \sqrt{4k_B TR \Delta\nu}$$

**Ideal diode equation:**

$$I(V) = I_0 [\exp(qV/kT) - 1]$$

**Dielectric polarization induced in a crystal:**

$$\mathbf{P}(t) = \varepsilon_0 \chi \mathbf{E}(t) + \varepsilon_0 \chi^{(2)} \mathbf{E}^2(t) + \varepsilon_0 \chi^{(3)} \mathbf{E}^3(t) + \dots$$

**Second-order polarization:**

$$P_i(\omega) = 2\varepsilon_0 \sum_{jk} \chi_{ijk}^{(2)} E_j(\omega) E_k(0)$$

**Thermal coefficient of resistance:**

$$\alpha(T) = \frac{1}{R} \left( \frac{dR}{dT} \right)$$

## Equation Summary

---

### Responsivity of a bolometer:

$$R_V = \frac{\alpha IR}{G_e(1 + i\omega C/G_e)}; \quad G_e = G - \alpha I^2 R$$

### Thermal conductance:

$$\tau = \frac{C}{G_e}$$

### Phonon noise:

$$NEP \approx \sqrt{4k_B T G}$$

### Helmholtz equation:

$$\nabla^2 E + k^2 E = 0$$

### Paraxial wave equation:

$$\partial^2 u / \partial x^2 + \partial^2 u / \partial y^2 - 2ik \partial u / \partial z = 0$$

### Complex beam parameter:

$$\frac{1}{q(z)} = \frac{1}{R(z)} - i \frac{\lambda}{\pi w^2(z)}$$

### Gaussian beam radius:

$$w(z) = \sqrt{\frac{-\lambda}{\pi \operatorname{Im}[1/q(z)]}}$$

$$w(z) = w_0 \left[ 1 + \left( \frac{\lambda z}{\pi w_0^2} \right)^2 \right]^{1/2}$$

$$w = w_0 \left[ 1 + \left( \frac{z}{z_c} \right)^2 \right]^{1/2}$$

### Gaussian beam phase radius of curvature:

$$R(z) = \left( \operatorname{Re} \left[ \frac{1}{q(z)} \right] \right)^{-1}$$

## Equation Summary

---

$$R(z) = z + \frac{1}{z} \left( \frac{\pi w_0^2}{\lambda} \right)^2$$

$$R = z + \frac{z_c^2}{z}$$

**Gaussian beam phase radius of curvature:**

$$\phi_0(z) \tan^{-1} \left( \frac{\lambda z}{\pi w_0^2} \right)$$

$$\phi_0 = \tan^{-1} \left( \frac{z}{z_c} \right)$$

**Confocal distance:**

$$z_c = \frac{\pi w_0^2}{\lambda}$$

**Gaussian beam:**

$$E(r, z) = E_0 \exp \left( -\frac{r^2}{w^2(z)} \right)$$

$$I(r, z) = I_0 \exp \left( -\frac{2r^2}{w^2(z)} \right)$$

**Total power in a Gaussian beam in a radius  $r$ :**

$$\frac{P(r)}{P_\infty} = 1 - \exp \left( -\frac{2r^2}{w^2(z)} \right)$$

**Full-width-half-maximum:**

$$FWHM = \sqrt{2 \ln(2)} w$$

## Equation Summary

---

### Edge taper:

$$T_e = \frac{P(a)}{P(0)}$$

$$T_e(\text{dB}) = -10 \log_{10}(T_e).$$

$$T_e(\text{dB}) = \frac{20}{\ln(10)} \left( \frac{a}{w} \right)^2 \approx 8.686 \left( \frac{a}{w} \right)^2$$

### Power lost due to truncation:

$$K_{lost}(r_t/w) = 1 - \exp(-2r_t^2/w^2) = 1 - T_e(r_t/w)$$

### Fraction of spillover power:

$$K_{spillover}(r_t/w) = \exp(-2r_t^2/w^2) = T_e(r_t/w)$$

### First dark ring of an Airy diffraction pattern:

$$\theta = 1.22\lambda/D_t$$

### Asymptotic beam growth angle:

$$\theta_0 = \tan^{-1} \left( \frac{\lambda}{\pi w_0} \right) \approx \frac{\lambda}{\pi w_0}$$

### ABCD matrices:

Propagation a distance  $d$ :  $\begin{bmatrix} 1 & d \\ 0 & 1 \end{bmatrix}$

Thin lens of focal length  $f$ :  $\begin{bmatrix} 1 & 0 \\ -1/f & 1 \end{bmatrix}$

Refraction at a curved dielectric interface:

$$\begin{bmatrix} 1 & 0 \\ \frac{(n_2 - n_1)}{R n_2} & \frac{n_1}{n_2} \end{bmatrix}$$

## Equation Summary

---

Reflection from a curved mirror:  $\begin{bmatrix} 1 & 0 \\ -2/R & 1 \end{bmatrix}$

$$q_{out} = \frac{A \cdot q_{in} + B}{C \cdot q_{in} + D},$$

$$\phi_{out} = \phi_{in} - \arg[A + B(1/q_{in})]$$

**Cylindrical Gaussian beam modes:**

$$\begin{aligned} \Psi_{pm}(r, \phi, z) = & \left[ \frac{2p!}{\pi(p+m)!} \right]^{0.5} \frac{1}{w(z)} \left[ \frac{\sqrt{2}r}{w(z)} \right]^m L_p^m \left( \frac{2r^2}{w^2(z)} \right) \\ & \cdot \exp \left[ \frac{-r^2}{w^2(z)} \right] \exp[-ikz] \exp \left[ -i \frac{\pi r^2}{\lambda R(z)} \right] \\ & \cdot \exp[-i(2p+m+1)\phi_0(z)] \exp(im\phi) \end{aligned}$$

**Rectangular Gaussian beam modes:**

$$\begin{aligned} \Psi_{mn}(x, y, z) = & \left[ \frac{1}{\pi w_x(z) w_y(z) 2^{m+n-1} m! n!} \right]^{0.5} \\ & \cdot H_m \left[ \frac{\sqrt{2}x}{w_x(z)} \right] H_n \left[ \frac{\sqrt{2}y}{w_y(z)} \right] \cdot \exp \left[ -\frac{x^2}{w_x^2(z)} - \frac{y^2}{w_y^2(z)} \right] \\ & \cdot \exp[-ikz] \cdot \exp \left[ -i \frac{\pi x^2}{\lambda R_x(z)} - i \frac{\pi y^2}{\lambda R_y(z)} \right] \\ & \cdot \exp \left[ -i \frac{(2m+1)}{2} \phi_{0x}(z) - i \frac{(2n+1)}{2} \phi_{0y}(z) \right] \end{aligned}$$

**Power coupling with relative displacement:**

$$K_{ab} = \frac{4}{\left( \frac{w_{0,b}}{w_{0,a}} + \frac{w_{0,a}}{w_{0,b}} \right)^2 + \left( \frac{\lambda \Delta z}{\pi w_{0,a} w_{0,b}} \right)^2}$$

**Power coupling with beam tilt:**

$$K_{ab} = \exp \left[ -(\pi w \theta / \lambda)^2 \right]$$

## Equation Summary

---

### Power coupling with lateral shift:

$$K_{ab} = \exp \left[ -(\Delta x/w)^2 \right]$$

### Scattering matrix:

$$\begin{pmatrix} \mathbf{b}_1 \\ \mathbf{b}_2 \end{pmatrix} = \begin{pmatrix} \mathbf{S}_{11} & \mathbf{S}_{12} \\ \mathbf{S}_{21} & \mathbf{S}_{22} \end{pmatrix} \begin{pmatrix} \mathbf{a}_1 \\ \mathbf{a}_2 \end{pmatrix}$$

### Cascading two scattering matrices:

$$\mathbf{S}^c = \begin{bmatrix} \mathbf{S}_{12}^a [\mathbf{I} - \mathbf{S}_{11}^b \mathbf{S}_{22}^a]^{-1} \mathbf{S}_{11}^b \mathbf{S}_{21}^a + \mathbf{S}_{11}^a & \mathbf{S}_{12}^a [\mathbf{I} - \mathbf{S}_{11}^b \mathbf{S}_{22}^a]^{-1} \mathbf{S}_{12}^b \\ \mathbf{S}_{21}^b [\mathbf{I} - \mathbf{S}_{22}^a \mathbf{S}_{11}^b]^{-1} \mathbf{S}_{21}^a & \mathbf{S}_{21}^b [\mathbf{I} - \mathbf{S}_{22}^a \mathbf{S}_{11}^b]^{-1} \mathbf{S}_{22}^a \mathbf{S}_{12}^b + \mathbf{S}_{22}^b \end{bmatrix}$$

### Transmission matrix:

$$\begin{pmatrix} \mathbf{b}_2 \\ \mathbf{a}_2 \end{pmatrix} = \begin{pmatrix} \mathbf{T}_{11} & \mathbf{T}_{12} \\ \mathbf{T}_{21} & \mathbf{T}_{22} \end{pmatrix} \begin{pmatrix} \mathbf{a}_1 \\ \mathbf{b}_1 \end{pmatrix},$$

$$\begin{pmatrix} \mathbf{T}_{11} & \mathbf{T}_{12} \\ \mathbf{T}_{21} & \mathbf{T}_{22} \end{pmatrix} = \begin{pmatrix} \mathbf{S}_{21} - \mathbf{S}_{22}[\mathbf{S}_{12}]^{-1} \mathbf{S}_{11} & \mathbf{S}_{22}[\mathbf{S}_{12}]^{-1} \\ -[\mathbf{S}_{12}]^{-1} \mathbf{S}_{11} & [\mathbf{S}_{12}]^{-1} \end{pmatrix}$$

### Cascading two transmission matrices:

$$\begin{aligned} \mathbf{T}^c &= \begin{bmatrix} \mathbf{T}_{11}^b & \mathbf{T}_{12}^b \\ \mathbf{T}_{21}^b & \mathbf{T}_{22}^b \end{bmatrix} \cdot \begin{bmatrix} \mathbf{T}_{11}^a & \mathbf{T}_{12}^a \\ \mathbf{T}_{21}^a & \mathbf{T}_{22}^a \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{T}_{11}^b \mathbf{T}_{11}^a + \mathbf{T}_{12}^b \mathbf{T}_{21}^a & \mathbf{T}_{11}^b \mathbf{T}_{12}^a + \mathbf{T}_{12}^b \mathbf{T}_{22}^a \\ \mathbf{T}_{21}^b \mathbf{T}_{11}^a + \mathbf{T}_{22}^b \mathbf{T}_{21}^a & \mathbf{T}_{21}^b \mathbf{T}_{12}^a + \mathbf{T}_{22}^b \mathbf{T}_{22}^a \end{bmatrix} \end{aligned}$$

### For an absorbing stop:

$$\mathbf{S} = \begin{bmatrix} \mathbf{S}_{11} & \mathbf{S}_{12} \\ \mathbf{S}_{21} & \mathbf{S}_{22} \end{bmatrix} = \begin{bmatrix} 0 & \mathbf{S}_{12} \\ \mathbf{S}_{21} & 0 \end{bmatrix}$$

$$S_{12}(\mu, \nu) = \int_{ap} [E_{\mu}^{-}(x', y')]^* E_{\nu}^{-}(x', y') dA,$$

$$S_{21}(\mu, \nu) = \int_{ap} [E_{\mu}^{+}(x', y')]^* E_{\nu}^{+}(x', y') dA$$



## Equation Summary

---

### For pure propagation:

$$\mathbf{S} = \begin{bmatrix} \mathbf{S}_{11} & \mathbf{S}_{12} \\ \mathbf{S}_{21} & \mathbf{S}_{22} \end{bmatrix} = \begin{bmatrix} 0 & \mathbf{D}_{21} \\ \mathbf{D}_{21} & 0 \end{bmatrix}$$

$$\mathbf{D}_{21}(m, n) = \delta_{mn} \exp[i\Delta\phi_m - ik(z_2 - z_1)]$$

$$\mathbf{T} = \begin{bmatrix} \mathbf{T}_{11} & \mathbf{T}_{12} \\ \mathbf{T}_{21} & \mathbf{T}_{22} \end{bmatrix} = \begin{bmatrix} \mathbf{D}_{21} & 0 \\ 0 & [\mathbf{D}_{21}]^{-1} \end{bmatrix} = \begin{bmatrix} \mathbf{D}_{21} & 0 \\ 0 & [\mathbf{D}_{21}]^* \end{bmatrix}$$

### Reflection at an interface:

$$\mathbf{S} = \begin{bmatrix} \mathbf{S}_{11} & \mathbf{S}_{12} \\ \mathbf{S}_{21} & \mathbf{S}_{22} \end{bmatrix} = \begin{bmatrix} \gamma \mathbf{L}_{11} & (1 + \gamma) \mathbf{I} \\ (1 + \gamma) \mathbf{I} & \gamma \mathbf{L}_{22} \end{bmatrix}$$

$$\gamma = \frac{1 - n}{1 + n}$$

### Ellipsoid of revolution:

$$a = (R_1 + R_2)/2, \quad b = \sqrt{R_1 R_2} \cos i$$

$$e = \sqrt{1 - b^2/a^2} = \sqrt{1 - 4R_1 R_2 \cos^2 i / (R_1 + R_2)^2},$$

$$x^2/a^2 + y^2/b^2 = 1$$

$$f = R_1 R_2 / (R_1 + R_2)$$

### Paraboloid of revolution:

$$y^2 + z^2 = 4px$$

$$p = R_1 \cos^2 i$$

$$x_C = R_1 \sin^2 i$$

$$y_C = R_1 \sin 2i$$

### Gaussicity of reflected beam:

$$K_G = 1 - w_m^2 \tan^2 i / (8f^2)$$

## Equation Summary

---

### Power scattered into the cross-polar direction:

$$K_{Xpol} = [w_m \tan i / (2f)]^2$$

### Polarization directions from a tilted grid:

$$\beta_{trans} = \pi/2 - \theta'$$

$$\beta_{refl} = \theta'$$

### Fraction of power transmitted from a tilted grid:

$$T_{trans} = \sin^2(\alpha - \theta')$$

$$T_{refl} = \cos^2(\alpha - \theta')$$

### Power transmission in an amplitude-splitting Michelson interferometer:

$$|T|^2 = 2|r|^2|t|^2[1 + \cos(k\Delta)]$$

### For a lossless system:

$$|r|^2 + |t|^2 = 1$$

$$|T|^2 = 2|r|^2(1 - |r|^2)[1 + \cos(k\Delta)]$$

### Power coupling in an interferometer:

$$K = |rt|^2 \left[ \frac{2 + \alpha^2 + 2[\cos k\Delta + \alpha \sin k\Delta]}{1 + \alpha^2} \right]$$

$$\alpha = \frac{\lambda\Delta}{2\pi w_0^2} = \frac{\Delta}{2z_c}$$

## Equation Summary

---

**Fractional power transmitted to the two output ports in a four-port interferometer:**

$$|T|^2 = [1 + \cos(k\Delta)]/2 \text{ and } |T|^2 = [1 - \cos(k\Delta)]/2$$

**Field from a corrugated conical horn:**

$$\mathbf{E}_{ap} = J_0 \left( 2.405 \frac{r}{a} \right) \exp \left( -ik \frac{r^2}{2L} \right) \mathbf{j}$$

**Phase error:**

$$s = d^2/8\lambda L$$

**Beam radius at the aperture of a conical corrugated horn:**

$$w_{ap} = 0.6435a$$

**Beam waist radius of a conical corrugated horn:**

$$w_0 = \frac{0.6435 a}{\sqrt{1 + \left( \frac{\pi (0.6435a)^2}{\lambda L} \right)^2}}$$

$$\Delta z = \frac{L}{1 + \left( \frac{\lambda L}{\pi (0.6435a)^2} \right)^2}$$

**Field from a pyramidal horn antenna:**

$$\mathbf{E}_{ap} = E_0 \cos \left( \frac{\pi x}{a} \right) \exp \left( -ik \left[ \frac{x^2}{2L_x} + \frac{y^2}{2L_y} \right] \right) \hat{\mathbf{j}}, \quad |x| \leq a/2, |y| \leq b/2$$

**Field from a diagonal horn antenna:**

$$\mathbf{E}_{ap} = \left[ \mathbf{i} \cos \left( \frac{\pi y}{a} \right) + \mathbf{j} \cos \left( \frac{\pi x}{a} \right) \right] \exp \left( -ik \left[ \frac{x^2 + y^2}{2L} \right] \right) \quad |x| \leq \frac{a}{2}; |y| \leq \frac{a}{2}$$

## Equation Summary

---

### Field from a smooth-walled conical horn antenna:

$$\mathbf{E}_{ap} \propto \{[J_0(\chi r/a) + J_2(\chi r/a) \cos 2\phi] \mathbf{i} + [J_2(\chi r/a) \sin 2\phi] \mathbf{j}\} \exp(-ikr^2/2L), \quad \chi = 1.841$$

### Profiled conical horn:

$$r(z) = r_{11} + (r_{12} - r_{11}) \times \left[ \frac{z}{L} (1 - A) + A \sin^2 \left( \frac{\pi}{2} \frac{z}{L} \right) \right]$$

### Collimating lens antenna:

$$d_{coll} = R \left( \frac{n}{n-1} \right)$$

### Hyperhemispherical lens antenna:

$$d_{hyper} = R \left( \frac{n+1}{n} \right)$$

### Critical angle for total internal reflection:

$$\theta_c = \sin^{-1}(1/n)$$

### Critical emission angle:

$$\phi_c = \sin^{-1} \left( \frac{n-1}{n} \right)$$

### Input and output waist positions:

$$d_{in} = f \pm M^{-1} (f^2 - f_0^2)^{0.5}$$

$$d_{out} = f \pm M (f^2 - f_0^2)^{0.5}$$

$$f_0 = \frac{\pi w_{0,in} w_{0,out}}{\lambda} \quad M = w_{0,out}/w_{0,in}$$

## Equation Summary

---

### Focal length:

$$f = \pm \frac{\left[ (M - M^{-1})^2 f_0^2 + d^2 \right]^{0.5} (M - M^{-1}) - 2d}{(M - M^{-1})^2}$$

### For $M=1$ :

$$f = \frac{d}{4} + \frac{f_0^2}{d}$$

## Bibliography

---

Boyd, R. W., *Nonlinear Optics*, Academic Press, Waltham, MA (2008).

Cunningham, P. D., Valdes, N. N., Vallejo, F. A., Hayden, L. M., Polishak B., Zhou, X.-H., Luo, J., Jen, A. K.-Y., Williams, J. C., and Twieg, R. J., “Broadband terahertz characterization of the refractive index and absorption of some important polymeric and organic electro-optic materials,” *J. Appl. Phys.* **109**, 043505, (2011).

Dexheimer, S., *Terahertz Spectroscopy Principles and Applications*, CRC Press, Boca Raton, FL (2008).

Ferguson, B. and Zhang, X.-C., “Materials for terahertz science and technology,” *Nature Mater.* **1**, 26–35 (2002).

Goldsmith, P. F., *Quasioptical Systems: Gaussian beam quasioptical propagation and applications*, Wiley–IEEE Press, Hoboken, NJ (1998).

Irwin, K. D. and Hilton, G. S., Transition-edge sensors, in *Cryogenic Particle Detection*, Ed. Enss, Topics Appl. Phys. vol. 99, 63–152, Springer–Verlag, Berlin (2005).

Lee, Y.-S., *Principles of Terahertz Science and Technology*, Springer, New York (2009).

Lesurf, J. C. G., *Millimetre-Wave Optics, Devices and Systems*, Taylor & Francis, London (1990).

Martin, D. H. and Bowen, J. W., “Log-wave optics,” *IEEE Trans. Microw. Theory Tech.* **41**, 10, 1676–1690 (1993).

Miles, R. E., Harrison, P., and Lippens, D., *Terahertz Sources and Systems*, Springer, New York (2001).

Milligan, T. A., *Modern Antenna Design*, Wiley–IEEE Press, Hoboken, NJ (2005).

Mittleman, D., Ed., *Sensing with Terahertz Radiation*, Springer Series in Optical Sciences **85** (2002).

Murphy, J. A., Egan, A., and Withington, S., “Truncation in millimeter and submillimeter-wave optical systems,”

## Bibliography

---

*IEEE Trans. Antennas and Propag.* **41**(10), 1408–1413 (1993).

Murphy, J. A., Trappe, N., and Withington, S., “Gaussian beam mode analysis of partial reflections in simple quasi-optical systems fed by horn antennas,” *Infrared Phys. Technol.* **44**, 289–297 (2003).

Naftaly, M. and Miles, R. E., “Terahertz time-domain spectroscopy for material characterization,” *Proc. IEEE* **95**(8), 1658–1665 (2007).

Olver, A. D., Clarricoats, P. J. B., Kishk, A. A., and Shafai, I., *Microwave Horns and Feeds*, IEEE Press, New York (1994).

O’Sullivan, C., Atad-Ettinger, E., Duncan, W., Henry, D., Jellema, W., Murphy, J. A., Trappe, N., van de Stadt, H., Withington, S., and Yassin, G., “Far-IR optics design and verification,” *Int. J. Infrared Milli.* **23**, 1029–1045 (2002).

Rieke, G. H., *Detection of Light: From the Ultraviolet to the Submillimeter*, Cambridge University Press, Cambridge (1994).

Richards, P. L., “Bolometers for infrared and millimeter waves,” *J. Appl. Phys.* **76**, 1–24 (1994).

Rogalski, A. and Sizov, F., “Terahertz detectors and focal plane arrays,” *Opto-Electron. Rev.* **19**, 346–404 (2011).

Rostami, A., Rasooli, H., and Baghban, H., *Terahertz Technology: Fundamentals and Applications*, Lecture Notes in Electrical Engineering vol. 77, Springer, Berlin (2011).

Sakai, K., Ed., *Terahertz Optoelectronics*, Topics in Applied Physics vol. 97, Springer, Berlin (2005).

Semenov, A. D., Gol’tsman, G. N., and Sobolewski, R., “Hot-electron effect in superconductors and its applications for radiation sensors,” *Supercond. Sci. Technol.* **15**, R1–R16 (2002).

## Bibliography

---

Siegel, P.H., "Terahertz technology," *IEEE Trans. Microw. Tech.* **50**, 910–928 (2002).

Siegman, A. E., *Lasers*, University Science Books, Sausalito, CA (1986).

Sizov, F., "THz radiation sensors," *Opto-Electron. Rev.* **18**, 10–36 (2010).

Sze, S. M. and Ng, K. K., *Physics of Semiconductor Devices*, John Wiley and Sons, Hoboken, NJ (2007).

van der Valk, N. C. J., Wenckebach, T., and Planken, P. C. M., "Full mathematical description of electro-optic detection in optically isotropic crystals," *J. Opt. Soc. Am. B* **21**, 3 (2004).

Van Rudd, J. and Mittleman, D. M., "Influence of substrate-lens design in terahertz time-domain spectroscopy," *J. Opt. Soc. Am. B*, **19**, 319–329 (2002).

Van Zyl, R., Perold, W., and Botha, R., The Gunn-diode: fundamentals and fabrication, *Proceedings of the 1998 South African Symposium on Communications and Signal Processing*, **COMSIG '98**, 407–412 (1998).

Zhang, X.-C., "Terahertz wave imaging: horizons and hurdles," *Phys. Med. Biol.* **47**, 3667–3677 (2002).

Zhang, X.-C. and Xu, J., *Introduction to THz Wave Photonics*, Springer, New York (2009).

Zmuidzinass, J. and Richards, P. L., "Superconducting detectors and mixers for millimeter and submillimeter astrophysics," *Proc. IEEE* **92**, 1597–1616 (2004).



## Index

---

- 1/ $f$  noise, 42
- ABCD* matrix, 61
- aberration, 72
- absorbance, 3
- absorption, 3
- active imaging, 97
- Airy pattern, 58
- amplitude distortion, 82
- asymptotic beam growth
  - angle, 60
- backward wave oscillator, 25
- beam modes, 74
  - see* Gaussian beam modes
- beam radius, 53, 60
- beam squint, 82
- beam waist, 54
  - radius, 54
- beamwidth,
  - see* half-power beamwidth
- Beer–Lambert law, 3
- bilayer, 44
- blackbody, 5
- bolometer, 41
- bound-to-continuum QCL
  - design, 21
- bremmstrahlung
  - radiation, 23
- chirped superlattice QCL
  - design, 21
- coherence length, 14
- coherent detection, 27
- complex radius of
  - curvature, 52, 61
- composite bolometers, 41
- confocal distance, 59, 60
- cross polarization, 82
- crystal detector,
  - see* rectifying diode
- cyclotron resonance
  - frequency, 22
- cyclotron resonance
  - masers, *see* gyrotron
- dark current, 32
- defocusing, 71
- detectivity, 30
- dielectric interface, 77
- difference frequency
  - generation, 12
- diffraction integrals, 96
- diffraction parameter, 86
- diffusion-cooled HEB, 48
- diplexer, 87
- direct detection, 27
- dispersion, 3
- edge taper, 56
- efficiency
  - aperture, 56
  - illumination, 56
  - power coupling, 69
  - spillover, 56
- electron acceleration, 6
- electron hopping, 41
- electro-optic crystal, 13
- electro-optic sampling, 39
- extrinsic detector, 32
- extrinsic excitation, 32
- femtosecond pulse, 7, 13, 17, 33, 39, 100
- free-electron laser (FEL), 24

## Index

- 
- frequency
    - down-conversion, 7
  - frequency multiplication, 7
  - frequency multipliers, 16
  - Fresnel number, 68
  - full-width-half-maximum (FWHM), 55
  - Gabor modes, 96
  - Gaussian beam modes, 51, 58
    - Gaussian–Hermite modes, 65
    - Gaussian–Laguerre modes, 64
    - high-order, cylindrical, 64
    - high-order, rectangular, 65
    - mode coefficients, 66
  - Gaussicity, 70
  - geometrical optics, 2
  - geometrical theory of diffraction, 96
  - GMBA, *see* Gaussian beam modes
  - Golay cell, 50
  - Gunn diode, 8
  - gyrotron, 22
  - half-power beamwidth, 55
  - HEB, *see* hot-electron bolometer
  - Helmholtz wave equation, 51
  - Hermite polynomials, 51
  - heterodyne receiver, 28, 37
  - horn antenna, 89
    - corrugated conical, 90
    - dual mode, 92
    - multimode, 93
    - shaped, 93
    - smooth-walled conical, 92
    - smooth-walled diagonal, 91
    - smooth-walled pyramidal, 91
  - hot-electron bolometer, 47
    - superconducting hot-electron bolometer, 48
  - hot spot, 48
  - idler beam, 15
  - illumination efficiency, *see* efficiency
  - impact ionization
    - avalanche transit time (IMPATT) diode, 9
  - incoherent detection, 27
  - interferometer
    - diffraction losses, 86
    - dual-beam, 85
    - Fabry–Pérot, 87
    - four-port dual-beam, 88
    - Martin–Puplett, 87
    - Michelson, 84, 87
  - intrinsic excitation, 32
  - Johnson noise, 31
  - Josephson effect, 45
  - Laguerre polynomials, 51

## Index

---

- Landau level, 20
- lasers, 6
- lens antenna, 94
- lenses
  - perfect, 76
- linac, 23
- linear electro-optic effect,
  - see* Pockels effect
- magnification, 95
- method of moments, 96
- microbolometer, 43
- mirrors
  - off-axis, 82
  - off-axis ellipsoid, 80
  - off-axis paraboloid, 81
- misalignment, 71
- mixer, 37
  - hot-electron bolometer, 48
- multimode horn, 42
- multiplexer, 87
- negative differential
  - resistance, 6
- noise equivalent power (NEP), 29, 30
- off-axis configuration, 80
- operator
  - linear scattering, 74
- optical heterodyning, 18
- optical parametric
  - oscillators, 15
- optical rectification, 12, 13
- optically pumped far-IR
  - gas laser, 19
- optically pumped THz
  - laser, *see* optically pumped far-IR gas laser
- parametric amplification, 15
- parametric generation, 15
- paraxial wave equation, 51
- partial coherence, 93
- partial reflections, 72, 76
- passive imaging, 97
- path length modulator, 84
- phase bunching, 22
- phase matching, 14, 15
- phase radius of curvature, 53
- phase slippage, 53, 61
- phonon-cooled HEB, 48
- phonon noise, 42
- photoconductive antenna (PCA), 17, 28, 33, 34
- photoconductive detector, 33
- photoconductive switch, 18
- photoconductors, 7
- photodetector, 27
- photomixer, 34
- photomixing, 7, 18, 28
- photon-assisted
  - tunneling, 38
- physical optics, 2, 96
- plane waves, 96
- Pockels effect, 27, 39
- polar crystal, 49
- polarizing grid, 83
- power coupling, 70
- probe pulse, 33
- propagation path
  - pure, 76

## Index

- 
- p-type germanium laser, 20
  - pump beam, 15
  - pyroelectric detector, 49
  - quantum cascade laser, 21
  - ray matrix, 61
  - ray tracing, 96
  - Rayleigh criterion, 58
  - Rayleigh range, *see* confocal distance
  - rectification, 27
  - rectifying diode, 35
  - reflectivity, 4
  - resonant phonon QCL design, 21
  - resonant tunnel diode, 11
  - responsivity, 29, 30, 42
    - of a bolometer, 42
  - roof mirror, 84
  - RTD, *see* resonant tunnel diode
  - scalar feed, *see* horn
    - antenna, corrugated conical
  - scattering matrix, 58, 72, 75
    - cascading of, 79
  - Schottky barrier diodes, 16
  - Schottky diode, 35, 37
  - second-harmonic generation, 12
  - shot noise, 31
  - sideband filter, 85
  - sidelobe, 89
  - signal beam, 15
  - signal-to-noise ratio, 29, 31
  - Smith–Purcell radiation, 26
  - SNR, *see* signal-to-noise ratio
  - spectroscopy, 99
    - time-domain, 100
  - spillover, 57
  - spillover efficiency, *see* efficiency
  - SQUID, *see* superconducting quantum interference device
  - standing wave, 72
    - antenna-fed systems, 78
  - standing waves, 78
  - storage ring, 23
  - sum-frequency generation, 12
  - superconducting quantum interference device (SQUID), 45
  - superconducting tunnel junctions, *see* superconductor–insulator–superconductor (SIS)
  - superconductor–insulator–superconductor (SIS), 38
  - superradiance, 26
  - susceptibility, 12
  - synchrotron, 23
  - synchrotron radiation, 23, 24
  - system design, 95

## Index

---

- T rays, 1
- taper efficiency,  
    *see* illumination efficiency
- thermal coefficient of  
    resistance, 42
- thermal detector, 27
- thermal noise, 31
- thermionic emission, 35
- THz band, 1
- THz gap, 2
- Ti:sapphire, 14, 17, 33, 39
- tolerancing, 71
- transferred-electron  
    device, 8
- transition-edge sensor, 44
- transition temperature,  
    44
- transmission, 3
- transmission matrices, 79
- transmission matrix, 73
- transmittance, 4
- truncation, 57, 72  
    aperture, 75
- tunable filters, 85
- tunnel injection transit  
    time (TUNNETT)  
    diode, 10
- ultrafast optics, 7
- undulator, *see* wiggler
- varactor diode, 16
- varistor diode, 16
- walk-off length, 14
- wiggler, 24



**Cr  idhe O'Sullivan** is a Senior Lecturer in the Department of Experimental Physics at the National University of Ireland (NUI), Maynooth. Her main research interests are in the fields of astronomy and terahertz optics and she works on a number of ongoing international research programs in the development of space- and ground-based astronomical instrumentation. Current projects include cosmic microwave background telescopes, bolometric interferometry and coupling schemes for detector arrays. Dr. O'Sullivan received a B.Sc. in Experimental Physics from University College Dublin in 1992 and a Ph.D. from the University of Cambridge in 1996. She previously worked on micromachined infrared bolometers in the Microelectronics Research Centre, University College Cork and on adaptive optics as a postdoctoral researcher in NUI, Galway. Dr. O'Sullivan has served on national committees for the Royal Irish Academy, the Institute of Physics, and SPIE's Terahertz Technology and Applications conference.



**J. Anthony Murphy** is Professor and Head of the Department of Experimental Physics at the National University of Ireland, Maynooth. His main research interests are in the area of far-infrared space optics and experimental cosmology, specifically, the cosmic microwave background. He was involved in the development of receiver systems for the European Space Agency Planck Surveyor Satellite and the Herschel Space Observatory. Prof. Murphy received his B.Sc. and M.Sc. in Experimental Physics in 1977 and 1979, respectively, from University College Cork. He subsequently obtained an M.S. in Physics in 1981, from the California Institute of Technology and a Ph.D. in

Physics in 1986 from University of Cambridge. From 1985 to 1987 he worked as a postdoctoral research associate at the Cavendish Laboratory Cambridge on receiver development for the James Clerk Maxwell Telescope (Hawaii). In 1988 he became a member of the lecturing staff of the Experimental Physics Department at NUI Maynooth. He is a member of the Institute of Electrical and Electronics Engineers and a Fellow of the Institute of Physics and the Royal Astronomical Society.

# Terahertz Sources, Detectors, and Optics

**Créidhe O'Sullivan and J. Anthony Murphy**

The region of the electromagnetic spectrum between microwaves and infrared radiation has come to be known as the “THz gap,” mainly due to the lack of readily available laboratory sources and detectors. Potential applications, however, particularly in medicine, security, and communications, have led to increased activity by the mainstream physics and engineering community in recent years.

The primary objective of this *Field Guide to Terahertz Sources, Detectors, and Optics* is to provide the reader with a concise description of the quasi-optical techniques used to design THz systems, as well as the basic principles of operation of the most common THz components in use today.

## SPIE Field Guides

The aim of each SPIE Field Guide is to distill a major field of optical science or technology into a handy desk or briefcase reference that provides basic, essential information about optical principles, techniques, or phenomena.

Written for you—the practicing engineer or scientist—each field guide includes the key definitions, equations, illustrations, application examples, design considerations, methods, and tips that you need in the lab and in the field.

John E. Greivenkamp  
Series Editor



**SPIE**

P.O. Box 10  
Bellingham, WA 98227-0010  
ISBN: 9780819491671  
SPIE Vol. No.: FG28



[www.spie.org/press/fieldguides](http://www.spie.org/press/fieldguides)

**SPIE**  
PRESS