1

Basic Optical Calculations

An excellent plumber is infinitely more admirable than an incompetent philosopher. The society which scorns excellence in plumbing because plumbing is a humble duty and tolerates shoddiness in philosophy because it is an exalted activity will have neither good plumbing nor good philosophy. Neither its pipes nor its theories will hold water.

Source: Gardner, J.W. 1961. *Excellence, Can We Be Equal and Excellent Too?* New York: Harper, p. 86.

1.1 Introduction

Okay, we've decided to build an electro-optical system. It's going to be so great that everybody who sees it will hate us. Now comes the fun part, the planning and designing, and then the hard part, the building and testing. To design and plan, we have to have some way of knowing how our brainchild ought to behave before it is built – that is, theory.

At the conceptual stage, the measurement principle is poorly understood, and many quite different schemes are suggesting themselves. To make sense of it, you need a white board and a couple of smart and experienced colleagues to bounce ideas off, plus some pointers on how to begin. The aim of this chapter is to equip you to do a conceptual instrument design on the back of an envelope, with numbers. It assumes some background in optics, especially some acquaintance with plane waves and Fourier transforms.[1]

The indispensable ingredients of a conceptual design are

- A measurement idea
- Operational requirements (field of view, scan speed, spot size, sensitivity, etc.)
- A photon budget
- A rough optical design
- A detection strategy
- A signal processing strategy
- The other kind of budget.

The best way to get them is through several sessions at that white board, with a lot of thought and calculation in between. (It is amazing how many people think they've finished with engineering calculations when they hand in their last exam, but that attitude is sudden death in the instrument-building business.) Once you have these, you can make a list of the main technical risks, in descending order of hairiness, and pretty soon you have a plan for how to proceed. The size of these technical risks is important – they can range from finding parts to violating laws of physics. Right at the beginning, we must decide whether the measurement is even possible, which requires more imagination than analysis. The invention of two-photon Doppler-free spectroscopy is a good example.

Example 1.1 ***Two-Photon Doppler-Free Spectroscopy***

Gas-phase spectroscopy[2] is limited at low pressure by the random thermal motions of the gas molecules, and at high pressures by their collisions. Molecules with different speeds along the laser beam axis experience different Doppler shifts,

1 Good books to fill in any needed background are listed in Appendix A.

2 This example is adapted from Vasilenko, L.S., Chebotaev, V.P., and Shishaev, A.V. (1970). *JETP Lett.* (English translation) 3: 161.

Building Electro-Optical Systems: Making It All Work, Third Edition. Philip C. D. Hobbs.

so their absorption features occur at different frequencies in the lab frame, leading to *Doppler broadening*. A two-photon transition involves the absorption of two photons, whose energies must sum to the transition energy. The absorption (and any resulting fluorescence) can be modulated by chopping the excitation beam. Using two excitation beams going in opposite directions, some events will involve absorption of one photon from each beam, which can occur only when both beams are unblocked by their choppers. If the modulation frequencies of the two beams are different, this part of the signal will appear at the sum and difference of the chopping frequencies. If a molecule has nonzero speed along the beam, then to leading order in V/c, it will see each beam shifted by

$$\Delta\nu_i = \frac{-\mathbf{k}_i \cdot \mathbf{v}}{2\pi}. \tag{1.1}$$

Since the two beams have oppositely directed $\mathbf{k}$ vectors, one will be upshifted and the other downshifted by the same amount; the sum of the photon energies, $h(\nu_1 + \nu_2)$ is unshifted. Thus, these mixing components are present only when the laser is tuned exactly to the rest-frame resonance frequency – they show no first-order Doppler broadening. An apparently inherent physical limit is circumvented with an extra chopper and a mirror or two; such is the power of a good measurement idea.[3]

Once the idea is in hand, analysis is needed to decide between competing alternatives. Such feasibility calculations are at the heart of electro-optical systems lore, because they help us develop a highly trained intuition; thus, we'll do lots of examples. The places where assumptions break down are usually the most instructive, so we'll also spend some time discussing some of the seedier areas of optical theory, in the hope of finding out where the unexamined assumptions lurk. We begin with wave propagation and imaging.

1.2 Wave Propagation

1.2.1 Maxwell's Equations and Plane Waves

Any self-respecting book on optics is obliged to include Maxwell's equations, which are the complete relativistic description of the electromagnetic field *in vacuo* and are the basis on which all optical instruments function. They are also useful for printing on T-shirts. Fortunately, this is a book of practical lore, and the full power of Maxwell's equations is almost never used in instrument design. The most basic equation that is generally useful is the *vector wave equation*,

$$\nabla^2\mathbf{E} - \frac{1}{c^2}\frac{\partial^2\mathbf{E}}{\partial t^2} = 0, \tag{1.2}$$

where c is the speed of light. Most of the time we will be calculating a monochromatic field, or at least a superposition of monochromatic fields. We can then separate out the time dependence as $\exp(-i\omega t)$ and write $k = \omega/c$, leaving the vector Helmholtz equation,

$$\left(\nabla^2 + k^2\right)\mathbf{E} = 0. \tag{1.3}$$

Its simplest solution is a vector plane wave,

$$\mathbf{E}(\mathbf{x}) = \mathbf{E}_0 e^{i\mathbf{k}\cdot\mathbf{x}}, \tag{1.4}$$

where the two fixed vectors are $\mathbf{E}_0$, the electric field vector, which may be complex, and $\mathbf{k}$ is the *wave vector*, whose magnitude $k = |\mathbf{k}| = \omega/c$ is called the *propagation constant*. If $\mathbf{E}_0$ is real, the field is said to be *linearly polarized* along $\mathbf{E}_0$; if its real and imaginary parts are the same size so that the instantaneous $\mathbf{E}$ rotates without changing length, the field is *circularly polarized*; otherwise, it's *elliptically polarized* (see Section 1.2.8. Of course, the actual instantaneous $\mathbf{E}$ is real-valued, but since this is always true, it's not necessary to carry along all the complex conjugate terms until we start doing nonlinear things, and leaving them out saves both time and blunders.[4]

Optics folk switch back and forth between $\mathbf{k}$ vectors and light rays, often without noticing what they're doing. This usually works fine, but it's liable to fall apart in situations where diffraction or birefringence are important, such as near a focus or a shadow boundary. We'll see much more of this in Section 1.3.5.

---

3 This idea has been extended to one-photon transitions by means of *saturation spectroscopy*, where the antipropagating beam is made strong enough to deplete the ground state population in the zero-velocity slice; this is an example of a *hole burning* measurement.

4 In signal processing, this is called the *analytic signal* convention, which we'll meet in Section 13.2.5.

Power flows parallel to $\mathbf{k}$ in an isotropic medium, but need not in an anisotropic one, so it is separately defined as the *Poynting vector* $\mathbf{S} = \mathbf{E} \times \mathbf{H}$ (see Sections 4.6.1 and 6.3.2). In the complex notation, the one-cycle average Poynting vector is $\mathbf{S} = \mathrm{Re}[\mathbf{E} \times \mathbf{H}^*]$.

1.2.2 Plane Waves in Material Media

So far, we have only considered propagation *in vacuo*. Electromagnetics in material media is enormously complicated on a microscopic scale, since there are $\sim 10^{22}$ scatterers/cm^3. Fortunately, for most purposes, their effects are well approximated by *mean field theory*, which smears out all those scatterers into a jelly that looks a lot like vacuum except for a change in the propagation velocity, the E/H ratio, and some loss. A plane wave entering a material medium via a plane surface remains a plane wave, with different $\mathbf{k}$ and $\mathbf{E}_0$.

In a medium, light travels at a speed $v = c/n$, where $1 \lesssim n \lesssim 4$. The constant n is the *refractive index* and is given by $n = \sqrt{\mu_r \varepsilon_r}$, where μ_r and ε_r are the relative magnetic permeability and dielectric constant[5] of the material at the optical frequency, respectively. Since μ_r is nearly always 1 in optics, $n = \sqrt{\varepsilon_r}$. In addition, the material will change the *wave impedance*,

$$Z = E/H = \sqrt{\mu/\varepsilon} = 4\pi \times 10^{-7} c/n \approx 377\ \Omega/n. \tag{1.5}$$

The analogy between wave impedance and transmission line impedance is a fruitful one.

In absorbing media, the refractive index is complex.[6] Assuming the medium is linear and time-invariant, the temporal frequency cannot change, so k has to be different in the medium; the new k is $k_n = nk_0$, where k_0 is the vacuum value. We usually drop the subscripts, so k is taken to be in the medium under consideration.

There are one or two fine points deserving attention here. One is that n varies with ω, a phenomenon known as *dispersion*. Besides making light of different colors behave differently, this leads to distortion of a time-modulated signal. The carrier wave propagates at the *phase velocity* $v_p = \omega/k$, but the envelope instead propagates, approximately unchanged in shape, at the *group velocity* v_g, given by

$$v_g = \partial\omega/\partial k. \tag{1.6}$$

Since the carrier propagates at the phase velocity v, as an optical pulse goes along its carrier "slips cycles" with respect to its envelope, that's worth remembering if you build interferometers. The group velocity approximation (1.6) holds for short propagation distances only, i.e. when the difference Δt in the transit times of different frequency components is much less than the pulse width τ. In the opposite limit, where $\Delta t \gg \tau$, the output is a time Fourier transform of the input pulse.[7]

The other fine point is that ε_r is in general a tensor quantity; there is no guarantee that the response of the material to an applied electric field is the same in all directions. In this book, we're concerned almost exclusively with the linear properties of optical systems, so a tensor is the most general local relationship possible. The resulting dependence of n on polarization leads to all sorts of interesting things, such as *birefringence* and *beam walkoff*. There are in addition strange effects such as *optical activity* (also known as *circular birefringence*), where the plane of polarization of a beam rotates slowly as it propagates. We'll talk more about these weird wonders in Chapters 4 and 6.

Aside: The Other Kind of Polarization

The dielectric constant ε_r expresses the readiness with which the charges in the medium respond to the applied electric field; it is given by $\varepsilon_r = 1 + 4\pi\chi$, where χ is the electric susceptibility (zero for vacuum); the electric polarization $\mathbf{P}$ is $\varepsilon_0 \chi \mathbf{E}$.

1.2.3 Phase Matching

The two basic properties of any wave field are amplitude and phase. At any point in space-time, a monochromatic wave has a unique phase, which is just a number specifying how many cycles have gone by since time $t = 0$. Since it's based on

5 Also called *electric permittivity*.

6 Complex refractive index $\tilde{n}$ is often quoted as $n + ik$, where n and k are real and positive, but it is conceptually simpler to leave $\tilde{n}$ complex, because the Fresnel formulae and Snell's law still work with absorbing media as long as they're expressed in terms of $\mathbf{k}$, which we'll do in Section 5.4.1.

7 The impulse response of a linearly dispersive medium is a chirp, and taking the Fourier transform is equivalent to convolving the given function with a chirp. This turns out to be important in digital signal processing (DSP), where it leads to the chirp-Z transform.

counting, phase is invariant to everything – it's a Lorentz scalar, i.e. it doesn't depend on your frame of reference or your coordinate system or anything else. This turns out to be a very powerful fact. The requirement for phase matching at optical interfaces is the physical basis of geometric optics.

Among solutions of the wave equation, a plane wave has the unique property of *translational invariance* – moving the origin of coordinates from 0 to $\mathbf{x}_0$ leaves everything the same except for an additive overall phase shift $\exp(\pm i\mathbf{k}\cdot\mathbf{x}_0)$, depending on your sign convention, equivalent to a pure time delay. A plane surface is also translationally invariant: it looks just the same from any point in the plane.

The electromagnetic field is continuous in source-free regions.[8] Thus at a planar interface, moving the reference point within the plane cannot change the relative phase between the fields on either side, because that would violate the continuity condition, so the additive phase has to be the same on both sides of the boundary. In terms of the $\mathbf{k}$ vectors, this means that the waves on either side must obey

$$\mathbf{k}_1 - \hat{n}(\hat{n}\cdot\mathbf{k}_1) = \mathbf{k}_2 - \hat{n}(\hat{n}\cdot\mathbf{k}_2) = \mathbf{k}_\parallel, \tag{1.7}$$

where $\mathbf{k}_1$ and $\mathbf{k}_2$ are the wave vectors in medium 1 and medium 2, respectively.

This *phase matching* condition is at the root of many important symmetry arguments in optics, such as grating diffraction, crystal optics, harmonic generation, acousto-optics, and simple refraction.

1.2.4 Refraction, Snell's Law, and the Fresnel Coefficients

If a plane wave encounters a plane interface between two semi-infinite slabs of index n_1 and n_2, as shown in Figure 1.1, the light is partially reflected and partially transmitted – a standard problem in undergraduate electromagnetics classes.

As we just discussed, the physics of the problem is phase matching. The propagation constant $k = 2\pi\nu/(cn)$ really is a constant in each medium, so for a given $k_\parallel$ there are two possible values of $k_\perp$ on each side, corresponding to waves propagating toward the interface or away:

$$k_\perp = \pm\sqrt{k^2 - k_\parallel^2} = \pm k\sin\theta. \tag{1.8}$$

Thus for a single source, we expect the fields to consist of an incident and a reflected plane wave on the input side and a single transmitted plane wave on the output side.[9] Forcing the tangential $\mathbf{k}$ vectors of all the waves be the same clearly gives the law of reflection for the reflected component, since only the $k_\perp$ component changes sign. From (1.8), it also gives *Snell's law* for the transmitted one: since the in-plane components have to match, the difference in k between media has to be taken up in the perpendicular component:

$$n_1\sin\theta_1 = n_2\sin\theta_2. \tag{1.9}$$

If there are m parallel planar interfaces, $k_\parallel$ is the same in all the layers, so since (in the jth layer) $k_j = n_j k_0$, we can use the phase matching condition to get $k_\perp$ in the jth layer:

$$k_\perp^2 = n_j^2 k_0^2 - k_\parallel^2. \tag{1.10}$$

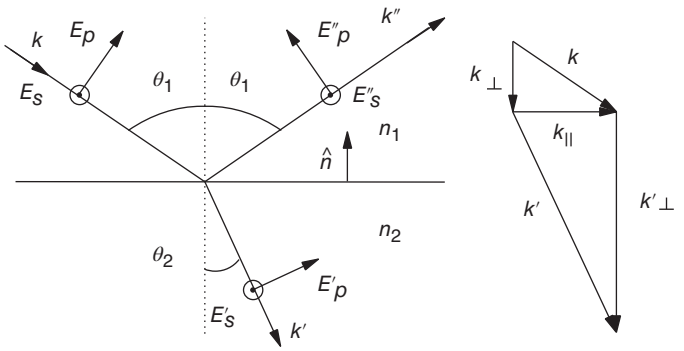


Figure 1.1 Refraction and reflection of a plane wave at a plane dielectric boundary. The angle of refraction θ_2 is given by Snell's law.

8 To be exact, at any boundary (real or notional) the tangential components of **E** and **H** are continuous, as are the perpendicular components of **D** and **B**.

9 This is of course a vector constraint – the tangential **k** vector has to lie in the plane of incidence as well as having the right length.

We'll see this again in Chapter 5's discussion of optical coatings. The continuity conditions on tangential **E** and perpendicular **D** across the boundary give the *Fresnel formulae* for the field amplitudes,

$$\frac{E_p''}{E_p} \equiv r_{p12} = -\frac{\tan(\theta_1 - \theta_2)}{\tan(\theta_1 + \theta_2)} = -\frac{n_2 \cos\theta_1 - n_1\sqrt{1 - \left[(n_1/n_2)\sin\theta_1\right]^2}}{n_2 \cos\theta_1 + n_1\sqrt{1 - \left[(n_1/n_2)\sin\theta_1\right]^2}}, \tag{1.11}$$

$$\frac{E_p'}{E_p} \equiv t_{p12} = \frac{2\sin\theta_1\cos\theta_2}{\sin(\theta_1 + \theta_2)} = \frac{2n_1\cos\theta_1}{n_2\cos\theta_1 + n_1\sqrt{1 - \left[(n_1/n_2)\sin\theta_1\right]^2}} \tag{1.12}$$

for light linearly polarized (i.e. **E** lying) in the *plane of incidence*. This plane is defined by the surface normal $\hat{\mathbf{n}}$ (unrelated to n) and $\mathbf{k}_{\text{inc}}$. In Figure 1.1 it is the plane of the page. For light linearly polarized perpendicular to the plane of incidence, these become

$$\frac{E_s''}{E_s} \equiv r_{s12} = -\frac{\sin(\theta_1 - \theta_2)}{\sin(\theta_1 + \theta_2)} = \frac{n_1 \cos\theta_1 - n_2\sqrt{1 - \left[(n_1/n_2)\sin\theta_1\right]^2}}{n_1 \cos\theta_1 + n_2\sqrt{1 - \left[(n_1/n_2)\sin\theta_1\right]^2}}, \tag{1.13}$$

$$\frac{E_s'}{E_s} \equiv t_{s12} = \frac{2\sin\theta_2\cos\theta_1}{\sin(\theta_1 + \theta_2)} = \frac{2n_1\cos\theta_1}{n_1\cos\theta_1 + n_2\sqrt{1 - \left[(n_1/n_2)\sin\theta_1\right]^2}}. \tag{1.14}$$

The two polarizations are known as p and s, respectively. As a mnemonic, s polarization means that **E** *sticks* out of the plane of incidence.[10] The quantities r and t are the reflection and transmission coefficient, respectively. These *Fresnel coefficients* act on the amplitudes of the fields.[11] The transmitted and reflected power ratios R and T, given by

$$\begin{aligned} R &= |r|^2 \\ T &= \frac{n_2\cos\theta_2}{n_1\cos\theta_1}|t|^2, \end{aligned} \tag{1.15}$$

are known as the *reflectance* and *transmittance*, respectively.

The Fresnel coefficients have fairly simple symmetry properties; if the wave going from n_1 to n_2 sees coefficients r_{12} and t_{12}, a wave coming in the opposite direction sees r_{21} and t_{21}, where

$$\begin{aligned} r_{p21} &= -r_{p12},\ t_{p21} = (n_1\cos\theta_1)/(n_2\cos\theta_2)t_{p12}, \\ r_{s21} &= -r_{s12},\ t_{s21} = (n_2\cos\theta_2)/(n_1\cos\theta_1)t_{s12}. \end{aligned} \tag{1.16}$$

The symmetry expressions for t_{21} are more complicated because they have to take account of energy conservation between the two media, where light travels at different speeds.[12]

1.2.5 Brewster's Angle

Especially sharp-eyed readers may have spotted the fact that if $\theta_1 + \theta_2 = \pi/2$, the denominator of (1.11) goes to infinity, so $r_p = 0$. At that angle, $\sin\theta_2 = \cos\theta_1$, so from Snell's law, $\tan\theta_1 = n_2/n_1$. This special value of θ_1 is called *Brewster's angle* θ_B. Note that the transmitted angle is $\pi/2 - \theta_B$, which is Brewster's angle for going from n_2 into n_1, as you can easily verify. If you've got an obnoxious reflection from a flat surface, and you can't get rid of it with a baffle or a drop of India ink at a focus, Brewster angle incidence with very pure p-polarized light is about the best existing technique (see Section 4.7.3).

Laser tube windows are always at Brewster's angle to reduce the round-trip loss through the cavity.[13] The loss in the s polarization due to four high-angle quartz-air surfaces is nearly 40% in each direction. Regeneration in the cavity greatly magnifies this gain difference, which is why the laser output is highly polarized. Brewster angle incidence is also commonly used in spectroscopic sample cells, Littrow prisms, and other high-accuracy applications using linearly polarized, collimated beams.

---

10 The s is actually short for *senkrecht*, which is German for perpendicular. The two polarizations are also called *transverse electric (TE)* and *transverse magnetic (TM)*, i.e. which field is sticking out of the plane of incidence. This nomenclature is more common in waveguide theory, where it means something different and makes much more sense.

11 There is another sign convention commonly used for the p-polarized case, where the incident and reflected **E** fields are taken in opposite directions, yielding a confusing sign change in (1.11). We adopt the one that makes $r_p = r_s$ at normal incidence.

12 See also Sections 1.2.2 and 5.4.1.

13 Except low-power He-Ne tubes, which have the mirrors fused to the ends.

Ideally, a smooth optical surface oriented at θ_B to an incoming p-polarized beam would reflect nothing at all, but this is not quite the case with real surfaces. Beams are never perfectly polarized, and all have some angular spread due to diffraction and wavefront error. Roughness and optical inhomogeneity make it impossible for every region to encounter the light beam at the same angle or with the same refractive index. Surface layers also prevent complete cancelling of the reflected wave, because their top and bottom reflections usually have different Brewster's angles and always have different phases due to the path delay between them. Below the critical angle, dielectric reflections always have phase 0 or π, so there's no way to tip the surface to zero out the phase-shifted reflection from a surface layer exactly.

Aside: Fossil Nomenclature
Back in 1808, when Malus discovered polarization by looking at reflections from a dielectric, he quite reasonably identified the *plane of polarization* with the plane of incidence. This conflicts with our modern tendency to fix on the **E** field in considering polarization. There are still some people who follow Malus's convention, so watch out when you read their papers. (The author has a Richardson Grating Laboratory catalog on his shelf that does this.)

1.2.6 Total Internal Reflection

If $n_1 > n_2$, there exists an angle θ_C, the *critical angle*, where Snell's law predicts that $\sin\theta_2 = 1$, so $\theta_2 = \pi/2$: grazing incidence. It is given by

$$\theta_C = \arcsin\left(n_2/n_1\right). \tag{1.17}$$

Beyond there, the surds in the Fresnel formulae (1.11)–(1.14) become imaginary, so t vanishes and r sits somewhere on the unit circle (the reflectivity is 1 and the elements of $\mathbf{E}'$ become complex).

This *total internal reflection* (TIR) is familiar to anyone who has examined a glass of water, or looked up from below the surface of a swimming pool. It is widely used in reflecting prisms and highway signs. There are two things to remember when using TIR: the reflection phase is massively polarization dependent and the fields extend beyond the surface, even though there is no propagating wave there. A TIR surface must thus be kept very clean, and at least a few wavelengths away from any other surface.

By putting another surface sufficiently close by, it is possible to couple light via the evanescent field, a phenomenon called *frustrated TIR* or more poetically, *evanescent coupling*. This is the optical analog of quantum mechanical tunneling.

The reflection phase represents a relative time delay of the propagating wave. The s-polarized wave is delayed more because it has a larger amplitude in the evanescent region, which requires more of a phase slip between the incident and reflected waves (remember the continuity conditions). This sort of physical reasoning is helpful in keeping sign conventions straight, although it is not infallible. The phase shift δ between s and p polarizations is[14]

$$\delta = \delta_s - \delta_p = -\arctan\frac{2\cos\theta_i\sqrt{\sin^2\theta_i - \left(n_2/n_1\right)^2}}{\sin^2\theta_i}. \tag{1.18}$$

1.2.7 Goos–Hänchen Shift

The angle-dependent phase shift on TIR functions much as dispersion does in the time domain, delaying different components differently. Dispersion causes the envelope of a pulse to propagate at the group velocity, which is different from the phase velocity. In the same way, the phase shift on TIR causes the envelope of a reflected beam to be shifted slightly in space from the incident beam, the *Goos–Hänchen shift*. It is less well known than the group velocity effect, mainly because the effect doesn't build up as it goes the way dispersion effects do, so the shift is small under normal circumstances, though large enough to cause image aberrations on reflection from TIR surfaces. The place it does become important is in multimode fiber optics, where a ray undergoes many, many reflections and the shift accordingly adds up.

The variation in the Goos–Hänchen shift comes from the speeding up of the wave that sticks out the most into the low-index material. It is responsible for the apparently paradoxical behavior of optical fiber modes near cutoff (see Section 8.3.1). We expect high-angle modes to slow down due to the decrease of k_z with angle – they spend more of their time bouncing back and forth instead of traveling down the fiber axis. In fact, they do slow down with increasing angle at

14 Born, M. and Wolf, E. (1983). *Principles of Optics*, 6e (corrected), 47–51. Pergamon: Oxford.

first, but then speed up again as they near cutoff, when the wave sticks farther and farther out into the low-index material. To leading order, the effect is the same as if the wave bounced off an imaginary surface one decay length into the low index material. Note that this does not contradict the last section; the phase shift is a delay, but the Goos–Hänchen shift makes the mode propagation anomalously fast.

1.2.8 Circular and Elliptical Polarization

What happens when the *p* and *s* components get out of phase with each other? The exponential notation, remember, is just a calculating convenience; real physical quantities always give real numbers. The instantaneous *E*-field strength is

$$E^2 = \left[\mathrm{Re}\left\{E_x e^{i\omega t}\right\}\right]^2 + \left[\mathrm{Re}\left\{E_y e^{i\omega t+\phi}\right\}\right]^2. \tag{1.19}$$

A linearly polarized monochromatic wave has an *E* that varies sinusoidally, passing through zero twice each cycle. When **E** has complex coefficients, the *p* and *s* components oscillate out of phase with one another. If the two are the same size and a quarter cycle apart, the real (i.e. physical) part of the **E** vector will spin through 2π once per cycle, without changing its length, like a screw thread. Its endpoint will traverse a circle so that this is known as *circular polarization*. Like screws, there is right and left circular polarization, but unlike screws, the names are backwards – left-circular polarization propagates with its **E** vector turning like a right-hand threaded bolt.

If the two components are not exactly equal in magnitude, or are not exactly $\pi/2$ radians apart, the vector will still rotate, but will change in length as it goes round, tracing an ellipse. This more general case is *elliptical polarization*. Circular and linear polarizations are special cases of elliptical polarization. Elliptical polarization can also be right or left handed.

1.2.9 Optical Loss

In a lossless medium, the **E** and **H** fields of a propagating wave are exactly in phase with each other. Any phase difference between them is due to absorption or gain in the material. A lossy material with dielectric constant $\varepsilon' - i\varepsilon''$ has a *loss tangent* $\delta = \varepsilon''/\varepsilon'$. In such a material, *H* lags *E* in phase by $1/2 \arctan \delta$.

1.3 Calculating Wave Propagation in Real Life

A real optical system is too complicated to be readily described in terms of vector fields, so being practical folk, we look for an appropriate sleazy approximation. We know that in a homogeneous and time-invariant medium, all propagating light waves can be decomposed into their plane wave spectra; at each (temporal) frequency, there will be a unique set of vector plane waves which combine to produce the observed field distributions. The approximations we will use are four:

1. *Scalar optics*. Replace vector field addition with scalar addition;
2. *Paraxial propagation*. Use an approximate propagator (the Huyghens integral) to calculate wave propagation, limiting us to beams of small cone angles;
3. *Fourier optics*. Use a simplistic model for how individual spatial Fourier components on a surface couple to the plane wave components of the incident light; and
4. *Ray optics*. Ignore diffraction by using an asymptotic theory valid as $\lambda \to 0$.

Analytical models, within their realm of applicability, are so much superior to numerical models in intuitive insight and predictive power that it is worth sacrificing quite a bit of accuracy to get one. Numerical models have their place, but the output is just a pile of special cases, so don't use them as a crutch to avoid hard thinking.[15]

1.3.1 Scalar Optics

If we shine an ideal laser beam (an apertured plane wave, i.e. perfectly collimated and having a rectangular amplitude profile) through a perfect lens and examine the resulting fields near focus, the result is a complete mess. There are nonzero field components along the propagation axis, odd wiggly phase shifts, and so on, due entirely to the wave and vector nature

15 This is not sour grapes. In 2006, the author spent some months writing a clusterized, optimizing 3D electromagnetic simulator for his antenna-coupled tunnel junction optical interconnect project, because there were none for sale. It's still actively maintained and in frequent use after almost 15 years. See https://electrooptical.net#POEMS.

of the fields themselves. The mess becomes worse very rapidly as the beam's half-angle θ approaches 90°. The effect is aggravated by the fact that no one really knows what a real lens does in sufficient detail to describe it accurately analytically – a real system is a real mess.[16]

For most optical systems, we don't have to worry about that, because empirically it doesn't affect real measurements much. Instead, we use *scalar optics*. Scalar optics is based on the replacement of the six components of the true vector electromagnetic field by a single number, usually thought of as being the electric field component along the (fixed) polarization axis. In an isotropic medium, the vector wave equation admits plane wave solutions whose electric, magnetic, and propagation vectors are constant in space, so that the field components can be considered separately, which leads to the *scalar Helmholtz equation*,

$$\left(\nabla^2 + k^2\right) E = 0. \tag{1.20}$$

(This rationale is nothing but a fig leaf, of course.) Any solution of (1.20) can be decomposed in a Fourier function space of plane waves, which are identified with one Cartesian component of the vector field. A scalar plane wave traveling along a direction $\mathbf{k}$ has the form

$$\psi(\mathbf{x}) = e^{i(\mathbf{k}\cdot\mathbf{x}-\omega t)}, \tag{1.21}$$

where the vector $\mathbf{k}$ has length $k = 2\pi/\lambda$. Conceptually, the true vector field can be built up from three sets of these.

The only difficulty with this in free space is that the field vectors of the plane waves are perpendicular to their propagation axes, so that for large numerical aperture,[17] where the propagation axes of the various components are very different, the vector addition of fields near a focus is not too well approximated by a scalar addition. Far from focus, this is not much of a worry, because the components separate spatially, as shown in Figure 1.2.

Aside: Plane Waves and δ-Functions

This separation is not entirely obvious from a plane wave viewpoint, but remember that plane waves are δ-functions in $\mathbf{k}$ space; that makes them just as singular in their way as δ-functions, so they aren't always an aid to intuition. Of course, it is not free space propagation that provides useful data, but the interaction of light with matter; boundaries of physical interest will in general mix the different polarization components. Such mathematical and practical objections are swept under the rug.[18]

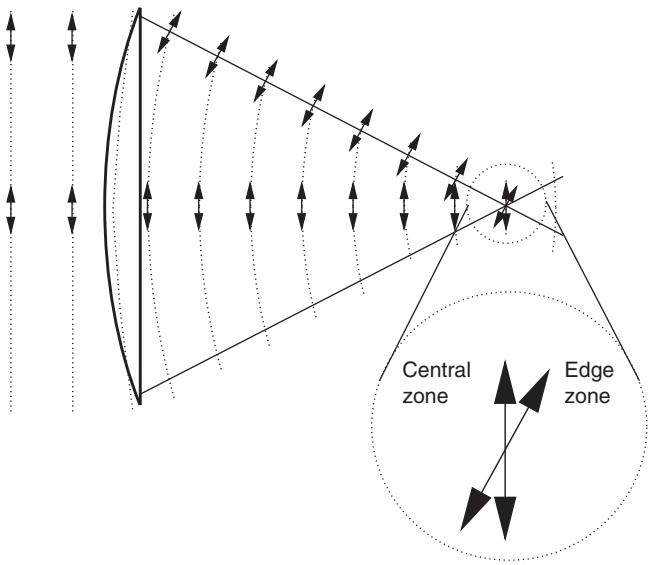


Figure 1.2 Scalar addition is a good approximation to vector addition except near high-NA foci or caustics, where light coming from very different directions gets concentrated.

16 Richards, B. and Wolf, E. (1959). *Proc. Roy. Soc.* (London) A253: 358, summarized in Stamnes, *Waves in Focal Regions* (see Appendix A).
17 The numerical aperture of a beam is given by NA $= n \sin\theta$, where n is the refractive index of the medium and θ is the half-angle of the beam cone. By Snell's law, the numerical aperture of a beam crossing an interface between two media at normal incidence remains the same.
18 The actual electromagnetic boundary conditions at surfaces of interest are very complicated, and usually poorly understood. Thus, most of the time the inaccuracies we commit by approximating the boundary conditions are smaller than our ignorance, alas.

Polarization and finite bandwidth effects are usually put in *by hand*. This means that we keep track of the polarization state of the beam separately and follow it through the various optical elements by bookkeeping; for frequency-dependent elements, we keep the frequency as an independent variable and integrate over it at the end. Such a procedure is inelegant and mathematically unjustified, but (as we shall see) it works well in practice, even in regimes such as high numerical aperture and highly asymmetric illumination, in which we would expect it to fail. Everyone in the optical systems design field uses it, and the newcomer would be well advised to follow this wholesome tradition unless driven from it by unusual requirements (and even then, to put up a fight).[19]

1.3.2 Paraxial Propagation

Discussions of beam propagation, pupil functions, optical and coherent transfer functions, and point spread functions take place with reference to the plane wave basis set. There is an infinite variety of such basis sets for decomposition of solutions of the scalar Helmholtz equation, nearly none of which are actually useful. Plane waves are one exception, and the Gauss–Legendre beams are another – or would be if they quite qualified. Gaussian beams (as they are usually called) don't even satisfy the true scalar Helmholtz equation because their phase fronts are paraboloidal rather than spherical and because they extend to infinity in both real and **k** (spatial frequency) space.

Instead, they satisfy the *slowly-varying envelope equation*, also known as the *paraxial wave equation*. First, we construct a field as a product of a plane wave $e^{i(kz-\omega t)}$ times an envelope function $\Theta(\mathbf{x})$ that varies slowly on the scale of a wavelength, i.e. $|(\nabla\Theta)/\Theta| \ll k$. We plug this field into the scalar Helmholtz equation, use the product rule for the Laplacian operator, and the subdominant term $d^2\Theta/dz^2$ is discarded, leaving a Schrödinger-type equation for the envelope Θ, the paraxial wave equation

$$\frac{d^2\Theta}{dx^2} + \frac{d^2\Theta}{dy^2} + 2ik\frac{d\Theta}{dz} = 0. \tag{1.22}$$

A general solution to this equation for all (*x,y,z*) is given by *Huyghens' integral*,[20]

$$\Theta(x,y,z) = -\frac{i}{\lambda}\iint_P \Theta(x',y',z')\frac{\exp\left[ik\frac{(x-x')^2+(y-y')^2}{2(z-z')}\right]}{(z-z')}dx'dy', \tag{1.23}$$

where P is the $x'y'$ plane. In diffraction theory (1.23) is also known as the *Fresnel approximation*. The Huyghens integral is an example of a *propagator*, an integral operator that uses the field values on a surface to predict those in the entire space. It is slightly inconvenient to lose the explicit phase dependence on z, but that can be recovered at the end by calculating the phase of an axial ray (one traveling right down the axis of the system) and adding it in. The Huyghens kernel depends only on $\mathbf{x} - \mathbf{x'}$ and so is a convolution (see Section 1.3.8), leading naturally to a Fourier space (**k** space) interpretation. In **k** space, (1.23) is

$$\Theta(x,y,z) = \iint_{P'} U(u,v)e^{i\frac{2\pi}{\lambda}(ux+vy)}e^{-i\frac{2\pi z}{\lambda}\frac{u^2+v^2}{2}}du\ dv, \tag{1.24}$$

where P' is the uv plane and U is the plane wave spectrum of Θ at $z = 0$, which is given by

$$U(u,v) = \iint_P \Theta(x,y,0)e^{-i\frac{2\pi}{\lambda}(ux+vy)}\frac{dx}{\lambda}\frac{dy}{\lambda} \tag{1.25}$$

The quantities u and v are the *direction cosines* in the x and y directions, respectively, and are related to the *spatial frequencies* k_x and k_y by the relations $u = k_x/k$, $v = k_y/k$. What we're doing here is taking the field apart into plane waves, propagating each wave through a distance z by multiplying by $\exp(ik_z z)$, and putting them back together to get the field distribution at the new plane. The (u,v) coordinates of each component describe its propagation direction. This is a perfectly general procedure in homogeneous media.

Aside: Use k Space

The real-space propagator (1.23) isn't too ugly, but the Rayleigh–Sommerfeld and Kirchhoff propagators we will exhibit in Section 9.3.2 are not easy to use in their real-space form. The procedure of splitting the field apart into plane waves,

19 For some of the reasons for this puzzling success, see Section 9.3.6.

20 It is sometimes stated that the Huyghens–Fresnel equation is a far-field approximation, amounting to Fraunhofer plus an additional quadratic term in the exponent, but this is not true. The validity of (1.22) only requires that the envelope Θ varies slowly on the scale of λ.

propagating them, and reassembling the new field is applicable to all these propagators, because in **k** space they differ only slightly (in their *obliquity factors*, of which more later). This is really the right way to go for hand calculations.

It is actually easier to spot what we're doing with the more complicated propagators because the $\exp(ik_z z)$ appears explicitly. The Huyghens propagator ignores the $\exp(ikz)$ factor and uses an approximation for $\exp[iz(k_z - k)]$, which obscures what's really happening.

In crystallography, **k** space is often called *reciprocal space*. That's a pretty good name because a lot of things are sort-of backwards: when n increases, light goes slower but **k** gets longer, and short times translate to high frequencies.

1.3.3 Gaussian Beams

The Gauss–Legendre beams are particular solutions to the paraxial wave equation. The general form of a zero-order Gauss–Legendre beam that travels along the z axis in the positive direction and whose phase fronts are planar at $z = 0$ is

$$\Phi(x, y, z, t) = \sqrt{\frac{2}{\pi}} \frac{1}{w(z)} \exp\left\{ i\phi(z) + \left(x^2 + y^2\right) \left[\frac{-1}{w^2(z)} + \frac{ik}{2R(z)} \right] \right\}, \tag{1.26}$$

where $R(z)$, $\phi(z)$, $w(z)$, z_R, and w_0 are given in Table 1.1. (Remember that the scalar field E is the envelope Φ multiplied by the plane wave "carrier" $e^{i(kz-\omega t)}$.) These parameters depend only on the beam waist radius w_0 and the wavelength λ of the light in the medium.

The envelope function is complex, which means that it modulates both the amplitude and the phase ϕ of the associated plane wave. This gives rise to the curved wavefronts (surfaces of constant phase) of focused beams, and also to the less well-known variations in $\partial\phi/\partial z$ with focal position, the Gouy phase shift.

Gaussian beams reproduce the ordinary properties of laser beams of small to moderate numerical aperture. They also form a complete set of basis functions, which means that any solution of (1.22) can be described as a sum of Gaussian beams. This useful property should not be allowed to go to the user's head; Gaussian beams have a well-defined axis, and so can only represent beams with the same axis. The number of terms required for a given accuracy and the size of the coefficients both explode as the beam axis departs from that of the eigenfunctions.

In this book, as in most of practical electro-optical instrument design, only this lowest-order mode, called TEM_{00}, is needed. This mode describes the field distribution of a good quality laser beam, such as that from a HeNe or circularized diode laser.

At large z, the Gaussian beam looks like a spherical wave with a Gaussian cutoff in u and v, but for small z, it appears to be a collimated beam. A reasonable limit to this region is where the beam area is twice that of the beam waist. The axial distance from the waist to this point is called the *Rayleigh range* z_R. It goes as $1/(\mathrm{NA})^2$ – the beam waist goes as $1/\mathrm{NA}$ and the angular width as NA. At $z = \pm z_R$, the $1/e^2$ beam diameter has increased by a factor of $\sqrt{2}$ so that the central intensity has halved.

The Gaussian beam is a paraxial animal: it's hard to make good ones of high NA. Its extreme smoothness makes it exquisitely sensitive to vignetting, which of course becomes inevitable as $\sin\theta$ approaches 1, and the slowly-varying envelope approximation itself breaks down as the numerical aperture increases (see Example 9.8).

There are a variety of parameters of Gaussian beams which are frequently of use, some of which are summarized in Table 1.1; P is the total power in watts, I is the intensity in $\mathrm{W/m^2}$, w is the $1/e^2$ intensity radius, w_0 is the beam waist radius, z_R is the Rayleigh range, and NA is measured at the $1/e^2$ intensity points in **k** space. Of interest in applications is the envelope phase, which shows a $\pm\pi/4$ phase shift (beyond the plane wave's $\exp(ikz)$) over the full depth of focus (twice the Rayleigh range), so that in a focused beam it is not a good assumption that the phase is simply $\exp(ikz)$ (see Example 9.4). This phase is exploited in phase contrast systems such as the Smartt interferometer of Section 12.7.6.

Aside: Gaussian Beams and Lenses

When a Gaussian beam passes through a lens, it is transformed into a different Gaussian beam. For the most part, ray optics is sufficient to predict the position of the beam waist and the numerical aperture, from which the waist radius and Rayleigh range can be predicted. There are some useful invariants of this process: for example, a Gaussian beam whose waist scans back and forth by b waist radii will be transformed into another beam whose waist scans b times the new waist radius. A position c times the Rayleigh range from the waist will image to a point c times the new Rayleigh range from the new waist. A corollary is that the number of *resolvable spots*, i.e. the scan range divided by the spot diameter, is also invariant. These invariants, which are not limited to the Gaussian case, allow one to juggle spot sizes, focal positions, and scan angles freely, without having to follow them laboriously through the whole optical system.

Table 1.1 TEM $_{00}$ Gaussian beam parameters (beam waist at $z = 0$).

Central intensity	$I_0 = 2P/(\pi w^2)$
Central intensity at the waist	$I_{0W} = 2P/(\pi w_0^2)$
Total power	$P = \pi w^2 I_0/2$
Beam waist radius (power density on axis $= I_0/e^2$)	$w_0 = \lambda/(\pi \mathrm{NA})$
$1/e^2$ power density radius	$w(z) = w_0[1 + (z/z_R)^2]^{1/2}$
3 dB power density radius vs. $1/e^2$ radius	$w_{1/2}(z) = 0.5887\, w(z)$
Radius within which $I >$ given I_{th}	$r = w/\sqrt{2\, \ln^{1/2}(I_0/I_{\mathrm{th}})}$
Power included inside $r \leq w$	86.4%
99% included power radius	$r_{99} = 1.517\, w$
Fourier transform pair	$\exp[-\pi(r/\lambda)^2] \supset \exp(-\pi \sin^2\theta)$
Separation of variables	$\exp(-\pi(r/\lambda)^2) = \exp[-\pi(x/\lambda)^2] \exp[-\pi(y/\lambda)^2]$
Numerical aperture ($1/e^2$ points in **k** space)	$\mathrm{NA} = \lambda/(\pi w_0)$
Radius of curvature of phase fronts	$R(z) = z + z_R^2/z$
Rayleigh range (axial intensity 50% of peak)	$z_R = \pi w_0^2/\lambda = \lambda/(\pi(\mathrm{NA})^2)$
Displacement of waist from geometric focus	$\Delta z \approx -z_R^2/f$
Envelope phase shift	$\phi(z) = \tan^{-1}(z/z_R)$
Equivalent projected solid angle	$\Omega'_{\mathrm{eq}} = \pi \mathrm{NA}^2 = \lambda^2/(\pi w_0^2)$

1.3.4 The Debye Approximation, Fresnel Zones, and The Fresnel Number

The plane wave decomposition of a given disturbance can be calculated from (1.25) or its higher-NA brethren in Section 9.3.6, and those work regardless of where we put the observation plane. When discussing the NA of a lens, however, we usually use a much simpler method: draw rays representing the edges of the beam and set $\mathrm{NA} = n \sin\theta$. This sensible approach, the *Debye approximation*, obviously requires the beam to be well represented by geometric optics, because otherwise we can't draw the rays – it breaks down if you put the aperture near a focus, for instance. We can crispen this up considerably via the *Fresnel construction*.

In a spherical wave, the surfaces of constant phase are equally spaced concentric hemispheres, so on a plane, the lines of constant phase are concentric circles, corresponding to annular cones, as shown in Figure 1.3. Drawing these circles at multiples of π radians divides the plane into annular regions of positive and negative field contributions, called *Fresnel zones*.

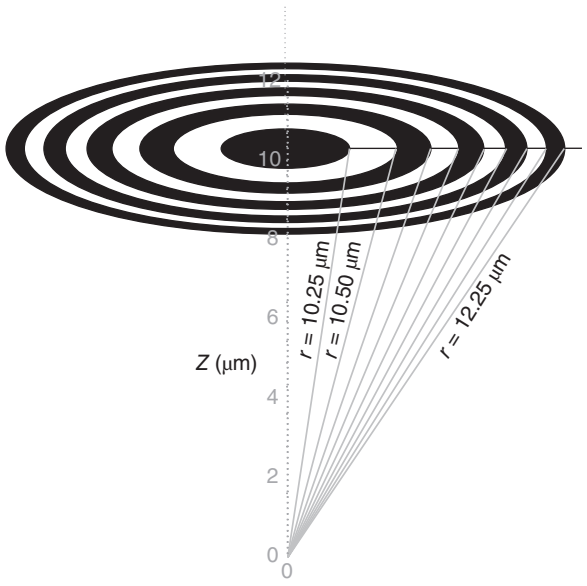


Figure 1.3 The Fresnel zone construction with $f = 10$ μm and $\lambda = 0.5$ μm. For a plane wave, taking the phase on axis as 0, alternating rings produce positive and negative field contributions at f, so blocking alternate ones (or inverting their phases with a $\lambda/2$ coating) produces a focus at f. For a converging spherical wave, all zones produce positive contributions at the focus.

The zones are not equally spaced; for a beam whose axis is along $\hat{z}$ and whose focus is at $z = 0$, the angular zone boundaries in the far field are at

$$\theta_n = \cos^{-1} \frac{1}{1 + (2n+1)\lambda/(4f)} \tag{1.27}$$

(the equation for the nth zone center is the same, except with $2n$ instead of $(2n + 1)$).

The *Fresnel number* N is the number of these zones that are illuminated. This number largely determines the character of the beam in the vicinity of the reference point – whether it is dominated by diffraction or by geometric optics. A collimated beam has $N = 1$, i.e. it is completely diffraction-dominated. The Debye approximation is valid in the geometric limit, i.e. $N \gg 1$.

Taking r to be the radius of the illuminated circle and applying a couple of trigonometric identities to (1.27) gets us a quadratic equation for N, the number of annular zone centers falling inside r. Assuming that $N\lambda \ll f$, this simplifies into

$$N = \frac{r^2}{\lambda z}, \tag{1.28}$$

which due to its simplicity is the usual definition of Fresnel number.

In small Fresnel-number situations, the focus is displaced toward the lens from its geometric position and diffraction is important everywhere, not just at the focus. The number of resolvable spots seen through an aperture of radius r is $N/2$.

Example 1.2 ***Gaussian Beams and Diffraction***

For small numerical apertures, the position of the beam waist does not coincide with the geometric focus, but is closer. This somewhat counterintuitive fact can be illustrated by considering a beam 40λ in radius, with a lens of $10^4\lambda$ focal length placed at its waist. The lens changes the Gaussian beam parameters, as we can calculate. Geometrically, incoming beam is collimated, so the focus is $10^4\lambda$ away, but in reality the Rayleigh range of the beam is $40(\pi/4)$ spot diameters, or 1260λ. This only 1/8 of the geometric focal length, so the lens makes only a small perturbation on the normal diffractive behavior of the original beam. At the geometric focus, $N = 40^2/10^4 = 0.16$, so the total phase change due to the lens is only $\pi/6$ across the beam waist.

1.3.5 Ray Optics

We all know that the way you check a board for warpage is by sighting along it, because light in a homogeneous medium travels in straight lines. The departures of light from straight-line propagation arise from nonuniformities in the medium (as in mirages) and from diffraction. Most of the time these are both small effects, so light can be well described by *rays*, thought of as vanishingly thin pencil beams whose position and direction are both well defined – the usual mild weirdness exhibited by asymptotic theories. Ray optics does not require the paraxial approximation, or even scalar waves.

In the absence of diffraction (i.e. as $\lambda \to 0$), the direction of propagation of a light beam in an isotropic medium is parallel to the gradient of the phase[21] $\nabla\phi$ (see Section 9.2.3). This means that a beam whose phase fronts are curved is either converging or diverging (see Section 9.2.2) and that rays can be identified with the normals to the phase fronts. Rays are the basis of elementary imaging calculations, as in the following example:

Example 1.3 ***Imaging with a Camera Lens***

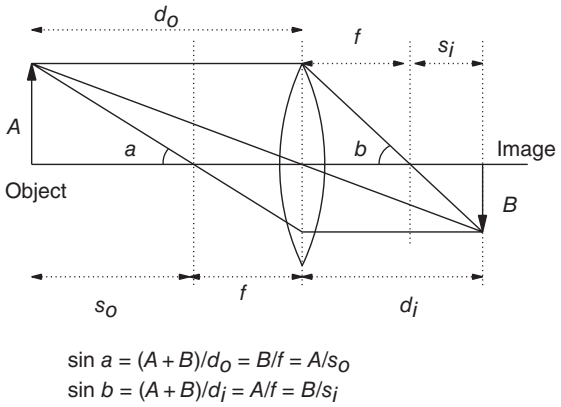
As a simple example of the use of ray optics, consider using a 35-mm camera to take an artistic head-and-shoulders portrait of a friend. For portraits, the most pleasing perspective occurs with a camera-to-subject distance of a few feet, 4 feet (1.3 m) being about optimal. What focal length lens is required?

The film frame is 24 by 36 mm in size and the outline of a human head and shoulders is about 400 by 500 mm. Thus, the desired magnification is 24/400, or 0.06. The rules of thin-lens optics are

1. Rays passing through the center of the lens are undeviated;
2. Rays entering parallel to the axis pass through the focus; and
3. The locus of ray bending is the plane of the center of the lens.

21 Born, M. and Wolf, E. (1983). *Principles of Optics*, 6e (corrected), p. 112. Pergamon: Oxford.

Figure 1.4 Portraiture with a 35-mm camera.



All of these rules can be fixed up for the thick-lens case (see Section 4.11.2). The similar triangles in Figure 1.4 show that the magnification M is

$$M = \frac{d_i}{d_o} \tag{1.29}$$

and elementary manipulation of the geometric identities shown yields

$$\frac{1}{d_o} + \frac{1}{d_i} = \frac{1}{f}, \tag{1.30}$$

$$s_o s_i = f^2. \tag{1.31}$$

Using (1.29) and (1.30), we find that

$$f = \frac{M d_o}{1 + M} \tag{1.32}$$

or 73.5 mm. Since the optimum distance is only a rough concept, we can say that a portrait lens for 35 mm photography should have a focal length of around 70–80 mm.

We can associate a phase with each ray, by calculating the phase shift of a plane wave traversing the same path. In doing this, we have to ignore surface curvature, in order that the wave remain plane. A shorthand for this is to add up the distances the ray traverses in each medium (e.g. air or glass) and multiply by the appropriate values of k.[22]

1.3.6 Lenses

In the wave picture, an ideal lens of focal length f transforms a plane wave $e^{ik(ux+vy)}$ into a converging spherical wave, whose center of curvature is at (uf, vf). It does so by inserting a spatially dependent phase delay, due to propagation through different thicknesses of glass. In the paraxial picture, this corresponds to a real-space multiplication by

$$L(x, y : f) = \exp\left[\frac{i\pi}{\lambda f}(x^2 + y^2)\right]. \tag{1.33}$$

Example 1.4 ***A Lens as a Fourier Transformer***
As an example of how to use the Huyghens propagator with lenses, consider a general field $\Theta(x, y, -f)$ a distance f behind a lens whose focal length is also f. The operators must be applied in the order the fields encounter them; here, the order is free-space propagation through f, followed by the lens's quadratic phase delay (1.33) and another free-space propagation through f. Using (1.23) twice, the field becomes

$$\Theta(x, y, +f) = \frac{i\lambda}{f}\int_{-\infty}^{\infty} d\frac{x'}{\lambda}\int_{-\infty}^{\infty} d\frac{y'}{\lambda} e^{-i\frac{2\pi}{\lambda f}(xx'+yy')}\Theta(x', y', -f) \tag{1.34}$$

22 This remains true even in anisotropic media, where double refraction occurs. It's **k** we're talking about, not the Poynting vector, despite what you'd see shining the beam on a white card – see Section 6.3.5).

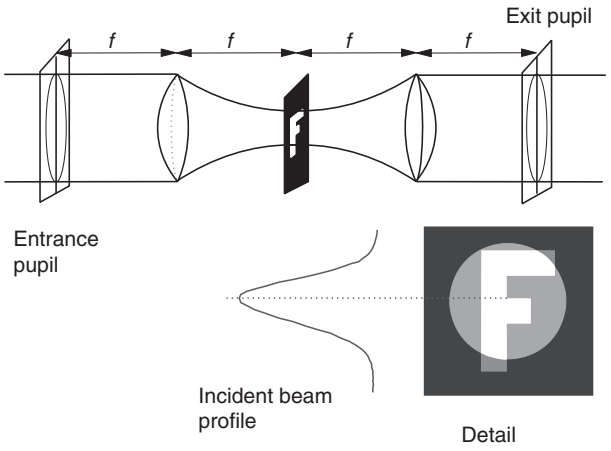


Figure 1.5 Spatial filtering.

which is a pure scaled Fourier transform. Thus, a lens performs a Fourier transform between two planes at $z = \pm f$. If we put two such lenses a distance $2f$ apart, as shown in Figure 1.5, then the fields at the input plane are reproduced at the output plane, with a Fourier transform plane in between. The image is inverted because we've applied two forward transforms instead of a forward $(-i)$ followed by a reverse $(+i)$ transform, so $(x, y) \rightarrow (-x, -y)$. This configuration is called a *4f relay* (see Section 9.3.12).

If we put some partially transmitting mask at the transform plane, we are blocking some Fourier components of Θ, while allowing others to pass. (A more general mask can apply attenuation and phase shifts as well.) Mathematically, we are multiplying the Fourier transform of Θ by the amplitude transmission coefficient of the mask, which is the same as convolving the image with the Fourier transform of the mask, appropriately scaled. This operation is *spatial filtering*, and is widely used.

The Fourier transforming property of lenses is extremely useful in both theory and applications. Perhaps surprisingly, it is not limited to the paraxial case, as we will see in Section 9.3.6.

1.3.7 Aperture, Field Angle, and Stops

Of the rays entering an optical system, some will make it out the other side and some won't. At most locations in a system, there is no sharp boundary between the two groups. At a given point in space, light going in some directions will make it and that going other directions will not; similarly, rays of a given angle in some places may make it and in other places not.

Each ray that fails to make it will wind up hitting some opaque surface. The surface that most limits the spatial field of a ray running parallel to the axis is called the *field stop* and that which most limits the angular acceptance of a point lying on the axis, the *aperture stop*. At these surfaces, the boundary between blocked and transmitted components is sharp. These locations are shown in Figure 1.6. It is common to put the aperture stop at a Fourier transform plane, since then all points on the object are viewed from the same range of angles. Optical systems image a volume into a volume, not just a plane into a plane, so the stops can't always be at the exact image and transform planes as these depend on the state of focus.

Aside: Vignetting

Aperture and field are defined in terms of axial points and axial rays. There's no guarantee that the aperture and field are exactly the same for other points and other directions. Rays that get occluded somewhere other than the field or aperture stops are said to have been *vignetted*. Vignetting isn't always bad; it's commonly used to get rid of badly aberrated rays that would degrade the image if they weren't intercepted. Better a dimmer image than a smeared-out one.

When a laser beam hits the edge of an aperture, this is also loosely termed vignetting, even when it does happen at one of the stops.

1.3.8 Fourier Transform Relations

Fourier transforms crop up all the time in imaging theory. They are a common source of frustration. We forget the transform of a common function, or can't figure out how to scale it correctly, and what was a tool becomes a roadblock. This is a pity

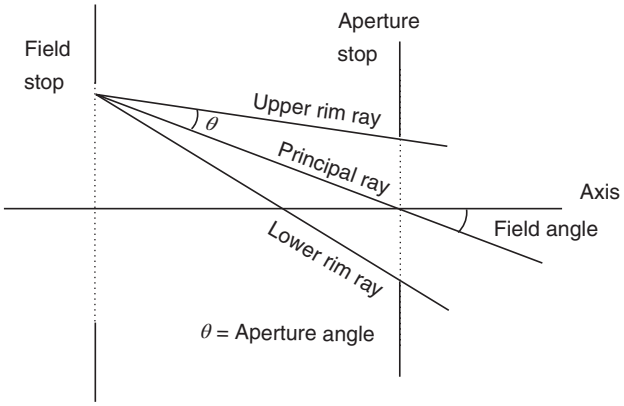


Figure 1.6 Definitions of aperture and field angle.

because Fourier transforms are so powerful and intuitive, once you have memorized a couple of basic facts and a few theorems. In an effort to reduce this confusion, here are a few things to remember. Following Bracewell, we put the factors of 2π in the exponents and write $g \supset G$ and $G = \mathscr{F}\, g$ for "g has transform G":

$$G(f) = \int_{-\infty}^{\infty} g(x) e^{-i2\pi fx} dx, \tag{1.35}$$

$$g(x) = \int_{-\infty}^{\infty} G(f) e^{i2\pi fx} df. \tag{1.36}$$

(It's customary to use lowercase in real space and uppercase in $\mathbf{k}$ space.) A side benefit of this is that we work exclusively in units of cycles.[23] One pitfall: for a wave traveling in the positive direction, the x and t terms in the exponent have opposite signs, so it's easy to get mixed up about forward and inverse transforms. Physicists and electrical engineers typically use opposite sign conventions (see Section 13.1).

Useful functions. The Heaviside unit step function $U(x)$ is 0 for $x < 0$ and 1 for $x > 0$. The derivative of $U(x)$ is the Dirac δ-function, $\delta(x)$. Sinc and jinc functions come up in connection with uniform beams: $\mathrm{sinc}x = \sin(\pi x)/(\pi x)$ and $\mathrm{jinc}x = J_1(2\pi x)/(\pi x)$. Even and odd impulse pairs $\mathbf{II}(x) = [\delta(x - 1/2) + \delta(x + 1/2)]/2$ and $\mathbf{I_I}(x) = [\delta(x + 1/2) - \delta(x - 1/2)]/2$ have transforms of $\cos(\pi x)$ and $i \sin(\pi x)$, respectively.

Conjugacy. Conjugate variables are those which appear multiplied together in the kernel of the Fourier transform, such as time in seconds and frequency in hertz. In optical Fourier transforms, the conjugate variables are x/λ and u, which as before is the direction cosine of the plane wave component on the given surface, i.e. $u = k_x/k$.[24]

Convolution. A convolution is the mathematical description of what a filter does in the real time or space domain, namely a moving average. If $g(x)$ is a given data stream and $h(x)$ is the impulse response of a filter (e.g. a Butterworth lowpass electrical filter, with x standing for time):

$$h(x)^* g(x) = \int_{-\infty}^{\infty} h(u) g(x - u) du = \int_{-\infty}^{\infty} g(u) h(x - u) du. \tag{1.37}$$

The second integral in (1.37) is obtained by the transformation $u \to x - u$; it shows that convolution is commutative: $g * h = h * g$. Convolution in the time domain is multiplication in the frequency domain,

$$\Phi(h * g) = HG. \tag{1.38}$$

This makes things clearer: since multiplication is commutative and associative, convolution must be too. A lot of imaging operations involve convolutions between a *point spread function* (impulse response) and the sample surface reflection coefficient (coherent case) or reflectance (incoherent case). The Huyghens propagator is also a convolution. Note that

23 In optical design, we use units of waves rather than cycles, but it's the same thing of course.
24 It's actually a bit more general; see Section 17.8.1.

one of the functions is flipped horizontally before the two are shifted, multiplied, and integrated. This apparently trivial point in fact has deep consequences for the phase information, as we'll see in a moment. The convolution theorem is also very useful for finding the transform of a function which looks like what you want, by cobbling together transforms that you know (see the example below).

Symmetry. By a change of variable in the Fourier integral, you can show that $g(-x) \supset G(-f), g^*(x) \supset G^*(-f), g^*(-x) \supset G^*(f)$, and (if g is real) $G(-f) = G^*(f)$.

Correlation and power spectrum. The cross-correlation g⋆h between functions g and h[25] is the convolution of $g(x)$ and $h^*(-x)$: $g \star h = g(x) * h^*(-x) \supset GH^*$. This can also be shown by a change of variables.

An important species of correlation is the *autocorrelation*, $g \star g$, whose transform is $GG^* = |G|^2$, the power spectrum. The autocorrelation always achieves its maximum value at zero (an elementary consequence of the Schwarz inequality) and all phase information about the Fourier components of g is lost.

Equivalent width. We often talk about the width of a function or its transform. There are lots of different widths in common use; 3 dB width, $1/e^2$ width, the Rayleigh criterion, and so on. When we come to make precise statements about the relative widths of functions and their transforms, we talk in terms of equivalent width or sometimes autocorrelation width. The equivalent width of a function is

$$w_e(g) = \frac{\int_{-\infty}^{\infty} g(x')dx'}{g(0)} = \frac{G(0)}{g(0)}. \tag{1.39}$$

It is obvious from this that if g and G are nonzero at the origin, the equivalent width of a function is the reciprocal of that of its transform. It is this relationship that allows us to say airily that an aperture 10λ wide has an angular spectrum 0.1 rad wide.[26]

Functions having most of their energy far from zero are not well described by an equivalent width. For example, if we move the same aperture out to $x = 200\lambda$, it will have a very large equivalent width (since $g(0)$ is very small), even though the aperture itself hasn't actually got any wider. Such a function is best described either by quoting its *autocorrelation width*, which is the equivalent width of the autocorrelation $g \star g$, or by shifting it to the origin. (We commonly remove the tilt from a measured wavefront, which is equivalent to a lateral shift of the focus to the origin.) Autocorrelations always achieve their maximum values at zero, because the contributions from all frequencies add up in phase there. Since the transform of the autocorrelation is the power spectrum, the autocorrelation width is the reciprocal of the equivalent width of the power spectrum.

Shifting. Given $g(x)$, shifting the function to the right by x_0 corresponds to subtracting x_0 from the argument. If $g(x)$ is represented as a sum of sinusoids, shifting it this way will phase shift a component at frequency f by fx_0 cycles:

$$\Phi(g(x - x_0)) = e^{-i2\pi f x_0} G(f). \tag{1.40}$$

Scaling. A feature 10λ wide has a transform 0.1 rad wide (in the small-angle approximation). Making it a times narrower in one dimension without changing its amplitude makes the transform a times wider in the same direction and a times smaller in height, without changing anything in the perpendicular direction:

$$g(at) \supset \frac{1}{|a|} G\left[\frac{f}{a}\right]. \tag{1.41}$$

You can check this by noting that the value of the transform at the origin is just the integral over all space of the function.

Integrating and differentiating. For differentiable functions, if $g \supset G$, then

$$\frac{dg}{dx} \supset i2\pi f G, \tag{1.42}$$

which is easily verified by integrating by parts. If g is absolutely integrable, then

$$\int^{x} g dx' \supset \frac{G}{i2\pi f} + K\delta(f), \tag{1.43}$$

where K is an arbitrary integration constant. It follows from (1.42) that the derivative of a convolution is given by

$$\frac{d}{dx}(h * g) = h * \frac{dg}{dx} = g * \frac{dh}{dx} \supset i2\pi f GH. \tag{1.44}$$

25 Note the difference in the symbols: ∗ is convolution, ⋆ is correlation. This isn't always obvious.

26 Things are done slightly differently in the circuits world, leading to mysterious factors of 2, e.g. a 1-s averaging window is said to have a bandwidth of 0.5 Hz. See Section 13.2.5 for the key to this mystery.

Power theorem. If we compute the central value of the cross-correlation of g and h, we get the odd-looking but very useful *power theorem*

$$\int_{-\infty}^{\infty} dx\, g(x)h^*(x) = \int_{-\infty}^{\infty} df\, G(f)H^*(f) \tag{1.45}$$

(think of g as voltage and h as current, for example). With the choice $g = h$, this becomes *Rayleigh's theorem*,[27]

$$\int_{-\infty}^{\infty} dx\, |g(x)|^2 = \int_{-\infty}^{\infty} df\, |G(f)|^2, \tag{1.46}$$

which says that the function and its transform have equal energy. This is physically obvious when it comes to lenses, of course.

Asymptotic behavior. Finite energy transforms have to fall off eventually at high frequencies, and it is useful to know how they behave as $f \to \infty$. A good rule of thumb is that if the nth derivative of the function leads to delta functions, the transform will die off as $1/f^n$. You can see this by repeatedly using the formula for the transform of a derivative until you reach delta functions, whose transforms are asymptotically constant in amplitude.

Transform pairs. Figure 1.7 is a short gallery of Fourier transform pairs.

Example 1.5 ***Cobbling Together Transforms***

In analyzing systems, we often need a function with certain given properties, but don't care too much about its exact identity as long as it is easy to work with and we don't have to work too hard to find its transform. For example, we might need a

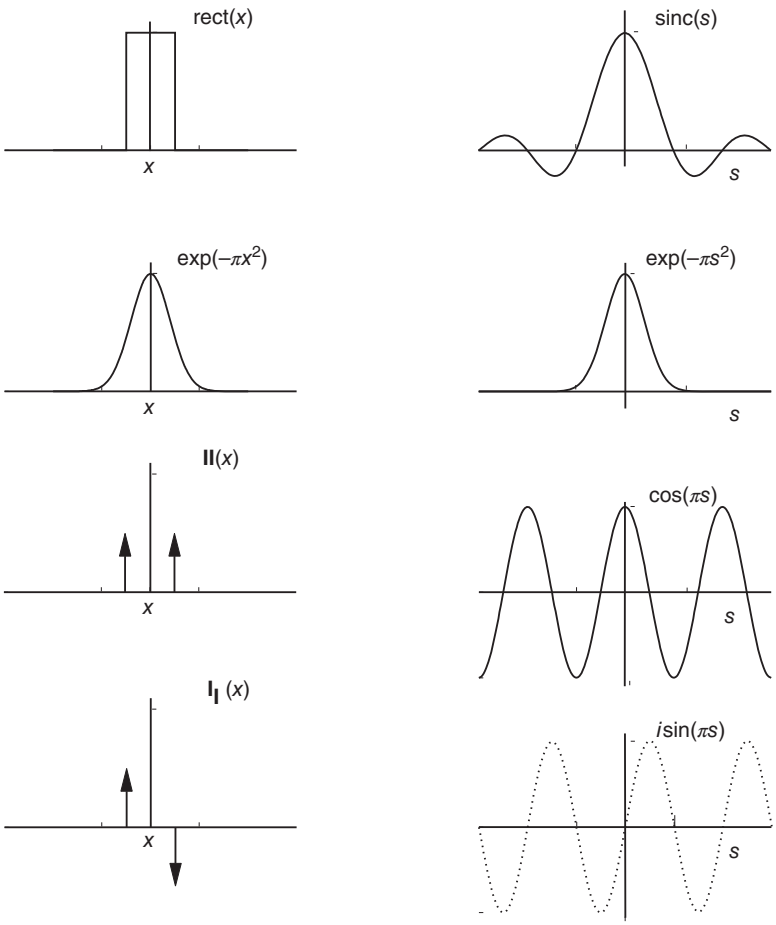


Figure 1.7 A pictorial gallery of Fourier transform pairs. Bracewell has lots more.

27 The same relationship in Fourier series is Parseval's theorem.

function with a flat top, decreasing smoothly to zero on both sides, to represent a time gating operation of width t_g followed by a filter. The gross behavior of the operation does not depend strongly on minor departures from ideal filtering, so it is reasonable to model this function as the convolution of $\text{rect}(t/t_g)$ with a Gaussian:

$$m(t) = \exp\left[-\pi(t/\tau)^2\right] * \text{rect}\left[\frac{t}{t_g}\right] \tag{1.47}$$

whose transform is

$$M(f) = \tau\ t_g e^{-\pi(f\tau)^2}\text{sinc}\left(f t_g\right). \tag{1.48}$$

One can write $m(t)$ as the difference of two error functions, but the nice algebraic properties of convolutions make the decomposed form (1.47) much more useful.

1.3.9 Fourier Imaging

We have seen that a lens performs a Fourier transform between its front and back focal planes, and that in **k** space, the propagation operator involves Fourier decomposing the beam, phase shifting the components, and reassembling them. There is thus a deep connection between the imaging action of lenses and the Fourier transform. Calculating the behavior of an imaging system is a matter of constructing an integral operator for the system by cascading a series of lenses and free space propagators, then simplifying. Nobody actually does it that way because it can easily run to 20th-order integrals. In a system without aberrations, we can just use ray optics to get the imaging properties, such as the focal position and numerical aperture, and then use at most three double integrals to get the actual fields, as in Example 1.4.[28]

Most of the time, we are discussing imaging of objects that are not self-luminous, so that they must be externally illuminated. Usually, we accept the restriction to *thin objects* – ones where multiple scattering can be ignored and the surface does not go in and out of focus with lateral position. The reasoning goes as follows: we assume that our incoming light has some simple form $E_{\text{in}}(x,y)$, such as a plane wave. We imagine that this plane wave encounters a surface that has an amplitude reflection coefficient $\rho(x,y)$, which may depend upon position, but not upon the angle of incidence, so that the outgoing wave is

$$E_{\text{out}}(x,y) = E_{\text{in}}(x,y)\rho(x,y) \tag{1.49}$$

and then we apply the Huyghens integral to E_{out} (or the Rayleigh–Sommerfeld integral if NA $\gtrsim 0.1$).

Small changes in height (within the depth of focus) are modeled as changes in the phase of the reflection coefficient. Since different plane wave components have different values of k_z, we apply a weighted average of the k_z values over the incident pupil function. The breakdown of this procedure due to the differences in $k_z z$ becoming comparable to a cycle gives rise to the limits of the depth of focus of the beam. We ignore the possibility that the height of the surface might be multiple-valued, e.g. a cliff or overhang, and any geometric shadowing.

A very convenient feature of this model, the one that gives it its name, is the simple way we can predict the angular spectrum of the scattered light from the Fourier transform of the sample's complex reflection $\rho(x,y)$. The outgoing wave in real space is the product ρE_{in}, so in Fourier space,

$$E_{\text{out}}(u,v) = E_{\text{in}}(u,v) * P(u,v), \tag{1.50}$$

where $\rho(x/\lambda, y/\lambda) \supset P$. The real power of this is that $E_{\text{out}}(u,v)$ is also the angular spectrum of the outgoing field so that we can predict the scattering behavior of a thin sample with any illumination we like.

If E_{in} is a plane wave of unit amplitude,

$$E_{\text{in}}(x,y) = \exp(i2\pi(u_{\text{in}}x + v_{\text{in}}y)/\lambda), \tag{1.51}$$

then its transform is very simple:

$$E_{\text{in}}(u,v) = \delta(u - u_{\text{in}})\delta(v - v_{\text{in}}). \tag{1.52}$$

By the sifting property of the delta function, the convolution of these two is

$$E_{\text{out}}(u,v) = P_{\text{in}}(u - u_{\text{in}}, v - v_{\text{in}}). \tag{1.53}$$

28 In *Modern Lens Design*, 2e, Smith gives a method for getting the third-order ray aberration coefficients from the thin-lens properties.

The angular spectrum of E_{out} is the spatial frequency spectrum of the sample, shifted by the spatial frequency of the illumination – the spatial frequencies add. In an imaging system, the spatial frequency cutoff occurs when an incoming wave with the largest positive u is scattered into the largest negative u the imaging lens can accept.[29] Since $u^2 + v^2 \leq \mathrm{NA}^2$, if the NAs of the illumination and the collecting lenses are equal, the highest spatial frequency an imaging system can accept is $2\ \mathrm{NA}/\lambda$.

The conceptual deficiencies of this procedure are considerable, even with thin objects. It works fine for large holes punched in a thin plane screen, but for more complicated objects such as transparent screens containing phase objects (e.g. microscope slides), screens with small features, or nearly anything viewed in reflection, the approximations become somewhat scalier. The conceptual problem arises right at the beginning, when we assume that we know *a priori* the outgoing field distributions at the boundary.

There is no real material which, even when uniform, really has reflection or transmission coefficients independent of angle and polarization at optical frequencies, and the situation is only made worse by material nonuniformity and topography. This and the scalar approximation[30] are the most problematic assumptions of Fourier optics; paraxial propagation is a convenience in calculations and not a fundamental limitation (see Section 9.3.5).

1.3.10 The Pupil

As anyone who has ever been frustrated by an out-of-focus movie knows, the image plane of an optical system is rather special and easily missed. Some other special places in an optical system are less well known. The most important of these is the *pupil*, which is an image of the aperture stop. If we look into the optical system from the object side, we see the *entrance pupil*. Looking from the image side, we see the *exit pupil*. By moving from side to side, we can locate the position of a pupil by how it moves in response, i.e. its parallax. (This is the same way we tell how far away anything appears.)

There's nothing magical about pupils, although they are talked about in terms that may confuse newcomers – they really are just places in an optical system that can be imaged where you want them and otherwise manipulated just as a focal plane can. Putting them in the right place will avoid pupil shift, which causes vignetting and its attendant loss of signal and resolution.

The aperture stop is usually put at a Fourier transform plane, to avoid nonuniform vignetting. The field distribution at a pupil then is the Fourier transform of that at the object or an image, appropriately scaled and with an obliquity correction. Equivalently, the field function at the transform plane is a scaled replica of the far-field diffraction pattern of the object, as derived using the Huyghens integral (1.23). When doing calculations, we generally use the normalized pupil coordinates (u, v), where $u = k_x/k$ and $v = k_y/k$, and sometimes also $w = k_z/k$. In an imaging system, the propagator is a convolution in real space. Since convolution in real space is equivalent to multiplication in $\mathbf{k}$ space, the imaging properties are controlled by the illumination pattern and detector sensitivity function at the transform plane. Since the transform plane is usually at the pupil, these are loosely called *pupil functions*, and are two-dimensional versions of the complex frequency response of an electronic system.[31] (They are still called pupil functions even when the transform plane is not at the pupil. Laziness is the father of invention.)

1.3.11 Pupil Problems

If the optical system has no moving parts, pupils behave just like other images. The same is true in postobjective scanning systems, where the scanner comes after the focusing lens; scanning makes the axis of the optical system pivot around the mirror's surface normal. Preobjective scanning is another matter. A preobjective scanner generally needs to be positioned at the pupil of the objective lens, and if you're using separate X and Y scan mirrors, they can't both be at the pupil. As we'll see in Section 7.10.7, this makes the pupil move around, which leads to vignetting and other unpleasantness.

Aside: Perspective

The center of the entrance pupil (or really, of the Fourier transform plane in the object space) is the *center of perspective*. If you're trying to make a panoramic view using an image mosaic, you'll want both foreground and background objects to have the same perspective – because otherwise, the positions of the joints between mosaic elements would have to be different depending on the distance. You can accomplish this by rotating the camera around its center of perspective.

29 There's nothing special about the choice of axes, so the limiting resolution might be different along y or at other angles.

30 There are edge diffraction effects that the scalar theory doesn't reproduce – see Section 9.2.7.

31 The analogy depends on the Debye approximation, so the exponential in/exponential out property of linear systems doesn't hold as accurately in Fourier optics as in most circuits, but it's still pretty good if the Fresnel number is high.

1.3.12 Connecting Wave and Ray Optics: *ABCD* Matrices

This section could be subtitled, "How to combine optical elements without drowning in multiple integrals." In an optical system consisting of lenses, mirrors, and free space propagation, it is possible to model the paraxial imaging properties by means of very simple transformation matrices, one for each element or air space, which are multiplied together to form a combined operator that models the entire system. Here we shall discuss the 2×2 case, appropriate for axially symmetric systems or for systems of cylindrical lenses whose axes are aligned. Generalization to 4×4 matrices is straightforward but more laborious.

In the small-angle approximation (where $\sin\theta \approx \theta$), a ray at height x above the optical axis and propagating at an angle θ measured counterclockwise from the optical axis is represented by a column vector $(x, \theta)^T$, and it transforms as

$$\begin{bmatrix} x \\ \theta \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} x' \\ \theta' \end{bmatrix}, \tag{1.54}$$

where the matrix *abcd* is the ordered product of the *ABCD* matrices of the individual elements. Let's do an example to see how this works.

Example 1.6 ***Deriving the* ABCD *Matrix for a Thin Lens***

In the ray tracing section, we saw that a thin lens brings all rays entering parallel to the axis through the focal point ($x = 0$, $z = f$), and that a ray passing through the center of the lens is undeviated. We can use these facts to derive the *ABCD* matrix for a thin lens, as shown in Figure 1.8. The undeviated central ray, $(0, \theta)^T$ is unchanged, so element B must be zero and element D must be 1. The ray parallel to the axis, $(1, 0)^T$, remains at the same height immediately following the lens so that element A is also 1. However, it is bent so as to cross the axis at f, so element C must be $-1/f$. Thus, a thin lens has an *ABCD* matrix given in Table 1.2. This procedure is easily extended to the 2D case (4×4 or 5×5 matrices).

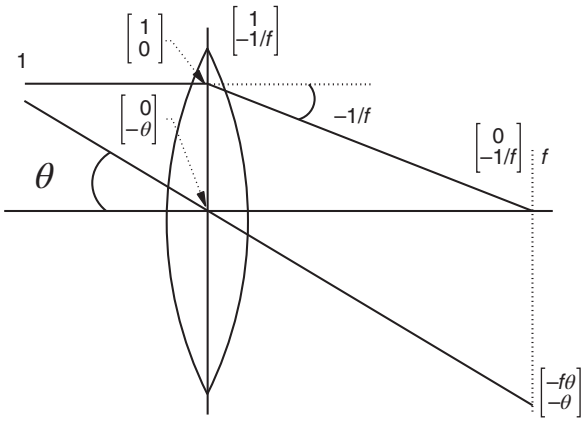
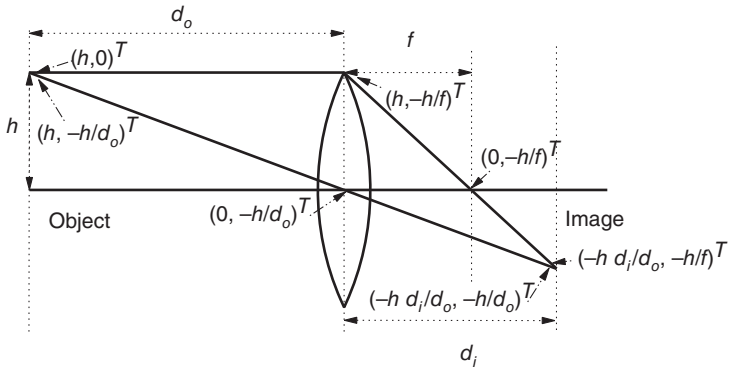


Figure 1.8 Action of a lens, for deriving its *ABCD* matrix.

Table 1.2 *ABCD* matrices for common operations.

Free space propagation through distance z	$\begin{bmatrix} 1 & z \\ 0 & 1 \end{bmatrix}$
Thin lens of focal length f	$\begin{bmatrix} 1 & 0 \\ -1/f & 1 \end{bmatrix}$
Magnification by M	$\begin{bmatrix} M & 0 \\ 0 & 1/M \end{bmatrix}$
Fourier transform	$\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$



$$\begin{bmatrix} 1 & d_i \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ -1/f & 1 \end{bmatrix} \begin{bmatrix} 1 & d_o \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1-d_i/f & d_o+d_i-d_o d_i/f \\ -1/f & 1-d_o/f \end{bmatrix}$$

$$= \begin{bmatrix} -d_i/d_o & 0 \\ -1/f & -d_o/d_i \end{bmatrix} = \begin{bmatrix} -M & 0 \\ -1/f & -1/M \end{bmatrix}$$

Figure 1.9 Imaging geometry with ray vectors and *ABCD* matrices: rays $(h,\theta)^T$ are successively multiplied by *ABCD* matrices corresponding to free space d_o, a lens of focal length f, and free space d_i. For the imaging condition, $1/d_o + 1/d_i = 1/f$, which makes the last two equalities true.

Optical layouts conventionally have light going from left to right[32], whereas matrix multiplication goes right to left. Thus, we have to write the matrix product backwards: the *ABCD* matrix of the first element encountered by the beam goes at the right, with subsequent operations left-multiplying it in succession, as in Figure 1.9.

It is straightforward to extend this formalism to small deviations from given angles of incidence, for example oblique reflection from a spherical mirror; when doing that, however, excellent drawings are required to avoid confusion about just what is going on.

Example 1.7 ***Portraiture Calculation Using* ABCD *Matrices***

Example 1.3 demonstrated how to find elementary imaging parameters such as magnification and focal length rapidly using the thin-lens rules on rays passing through the center of the lens and rays passing through the focus. Let us follow the path of a more general paraxial ray using *ABCD* matrices. We note first that the light from the object propagates through $d_o = 1300$ mm of free space, then a thin lens of focal length $f = 73.5$ mm, and finally another free space propagation through a distance d_i. The column vector representing the ray must be acted on by the matrix operators (written in reverse order as already noted):

$$\begin{aligned} \begin{bmatrix} x' \\ \theta' \end{bmatrix} &= \begin{bmatrix} 1 & d_i \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ -1/f & 1 \end{bmatrix} \begin{bmatrix} 1 & d_o \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ \theta \end{bmatrix} \\ &= \begin{bmatrix} 1-\frac{d_i}{f} & d_o+d_i-\frac{d_i d_o}{f} \\ -\frac{1}{f} & 1-\frac{d_o}{f} \end{bmatrix} \begin{bmatrix} x \\ \theta \end{bmatrix} \\ &= \begin{bmatrix} 1-0.0136 d_i & 1300+18.68 d_i \\ -0.0136 & -16.68 \end{bmatrix} \begin{bmatrix} x \\ \theta \end{bmatrix}. \end{aligned} \tag{1.55}$$

Comparing the algebraic form of the matrix product in (1.55) to the prototypes in Table 1.2, it is apparent that the combination of a lens plus free space on either side behaves as a Fourier transformer (scaled by a magnification of $-f$) when $d_o = d_i = f$. Furthermore, the imaging condition demands that all rays leaving an object point coincide at the same image point; this means that b, the (1, 2) element of the matrix, must be zero, which reproduces (1.30). These sorts of considerations are very valuable for more complex systems, where the thin lens ray picture is cumbersome.

32 See Section 4.11.

1.3.13 Extended *ABCD* Matrices

Useful as these matrices are, they can't model such elementary operations as the addition of a thin prism of angle ϕ and index n, which requires adding an angle $\Delta\theta = (n-1)\phi$, or a decentered lens, which is equivalent to a lens with a shift of δx on one side and $-\delta x$ on the other. This can be patched up by using the augmented vectors $[x,\ \theta,\ 1]^T$ and 3×3 matrices. A general element producing a transformation *ABCD* and then adding an offset $[\Delta x, \Delta\theta]^T$ is then

$$\begin{bmatrix} a & b & \Delta x \\ c & d & \Delta\theta \\ 0 & 0 & 1 \end{bmatrix}. \tag{1.56}$$

This is especially useful in getting some idea of the allowable tolerances for wedge angle, tilt, and decentration; a lens of focal length f, decentered by a distance d, adds an angular offset $\Delta\theta = d/f$. Similarly, a window of thickness t and index n, whose axis is α degrees off the normal, looks like a free space propagation of t/n with a spatial offset of $\Delta x = \alpha t(n-1)$. The author often uses this method generalized to 2D (with 5×5 matrices) for rough optical layouts.

1.3.14 *ABCD* Matrices and Wave Optics

These matrix operators ignore wave effects and are completely unable to cope with absorbing or partially scattering objects such as beamsplitters and diffraction gratings. While these can, of course, be put in by hand, a more general operator algebra is desirable, which would take account of the wave nature of the light and model light beams more faithfully.

Nazarathy and Shamir[33] have produced a suitable operator algebra for Fourier optics. The key simplification in use is that they have published a multiplication table for these operators, which allow easy algebraic simplification of what are otherwise horrible high-order multiple integrals. These transformations are in principle easy to automate and could be packaged as an add-on to symbolic math packages. This algebra takes advantage of the fact that commonly encountered objects such as lenses, gratings, mirrors, prisms, and transparencies can be modeled as operator multiplication in the complex field picture, which (as we have seen above) doesn't work in the ray picture.

Another way of coming at this is to use *ABCD* matrices for the operator algebra and then convert the final result to a Huyghens integral. In the paraxial picture, an axisymmetric, unaberrated, unvignetted optical system consisting of lenses and free space can be expressed as a single *ABCD* matrix, and any *ABCD* matrix with $d \neq 0$ can be decomposed into a magnification followed by a lens followed by free space:

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} 1 & z \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ -\frac{1}{f} & 1 \end{bmatrix} \begin{bmatrix} M & 0 \\ 0 & \frac{1}{M} \end{bmatrix}, \tag{1.57}$$

$$\text{where } M = \frac{1}{d}, \quad z = \frac{b}{d}, \quad f = -\frac{1}{cd}. \tag{1.58}$$

Element a does not appear because that degree of freedom is used up to ensure that the determinant of the matrix is unity, as required by the conservation of phase space volume.[34] A magnification by M corresponds to the integral operator

$$\Theta(x,y) = \frac{1}{M}\iint dx'dy'\Theta(x',y')\delta\left[x' - \frac{x}{M}\right]\delta\left[y' - \frac{y}{M}\right]. \tag{1.59}$$

Identifying these matrix operators with the corresponding paraxial integral operators (1.23), (1.33), and (1.59), we can construct the equivalent integral operator to a general *ABCD* matrix with $d \neq 0$:

$$\Theta(x,y) = \frac{-i}{zM\lambda}\iint dx'dy'\Theta\left[\frac{x'}{M}, \frac{y'}{M}\right]\exp\left[\frac{i\pi}{\lambda}\left(\frac{(x-x')^2+(y-y')^2}{z} + \frac{x'^2+y'^2}{f}\right)\right]. \tag{1.60}$$

This transformation is simple, and it can save a *lot* of ugly integrals. The special case where $d = 0$ corresponds to a lens, followed by a scaled Fourier transform:

$$\begin{bmatrix} a & b \\ -1/b & 0 \end{bmatrix} = \begin{bmatrix} M & 0 \\ 0 & 1/M \end{bmatrix} \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ -1/f & 1 \end{bmatrix}. \tag{1.61}$$

33 Nazarathy, M. and Shamir, J. (1982). First-order optics—a canonical operator representing lossless systems, *J. Opt. Soc. Am.* 72: 356–364.
34 AKA conservation of radiance, conservation of étendue, or the Lagrange invariant – see Section 9.2.9.

The constraint that $c = -1/b$ keeps the determinant 1, as before. Here the parameters are $M = -b, f = -b/a$ so that the equivalent integral in the wave picture is

$$\Theta(x,y) = \frac{-iM}{\lambda} \iint dx'dy'\Theta(Mx', My') \exp\left[\frac{i\pi M}{\lambda}\left(xx' + yy' + \frac{x'^2 + y'^2}{f}\right)\right]. \tag{1.62}$$

These two equivalences allow wave and ray descriptions of the same optical system to be freely interchanged, which is very convenient in calculations.

In the integral representation, offsets are modeled as convolutions with shifted δ-functions; a shift of ξ is a convolution with $\delta(x - \xi)$.

Aside: Complex *ABCD* Matrices and Diffraction

Siegman[35] shows that a Gaussian amplitude apodization can be modeled using the *ABCD* matrix for a thin lens with an imaginary focal length. This isn't magic, It's just that a thin lens is a multiplication by an imaginary parabolic exponential, $I(x) = \exp[i\pi x^2/(\lambda f)$. A Gaussian of $1/e^2$ radius $w, A(x) = \exp(-x^2/w^2)$, so mathematically speaking it's a lens of focal length $i\pi w^2/\lambda$.[36] Thus, by making the first (rightmost) *ABCD* matrix a Gaussian aperture,

$$\begin{bmatrix} 1 & 0 \\ -i\lambda/(\pi w^2) & 1 \end{bmatrix}, \tag{1.63}$$

you can carry the beam radius right through the *ABCD* calculation, including converting it to a Helmholtz integral. By keeping track of the diffraction effects automatically, this makes it simple to find the true beam waist, for instance, and if you're building interferometers with very small diameter beams, allows you to calculate the phase front matching at the beam combiner. It's pretty slick.

1.3.15 Source Angular Distribution: Isotropic and Lambertian Sources

A light source (such as the quiet Sun) whose output is independent of direction is said to be *isotropic*; there's no special direction. Light inside an integrating sphere (Section 5.8.7) is also isotropically distributed because there's no special direction. When there's a surface involved, though, things change on account of obliquity. If you shine a light on a perfectly matte-finished surface, its surface looks equally bright no matter what angle you look from. If you tilt it, it gets foreshortened by the perspective, but if you looked at it from a distance through a drinking straw, you wouldn't be able to tell from its brightness whether it was tilted or not. A surface that passes the drinking-straw test is said to be *Lambertian*, and the scattering is said to be *diffuse*.

If you replace your eye with a photodiode, each drinking-straw patch of emitting or detecting surface contributes the same amount of photocurrent. As the angle increases, the patches get longer like evening shadows, so $\cos\theta$ times fewer patches will fit on the surface. Another way to put this is that the total projected area goes down like the cosine of the angle of incidence, so the detected photocurrent from unit surface area will be multiplied by the *obliquity factor* $\cos\theta$. Obliquity factors come in just about everywhere – often disguised as $\hat{\boldsymbol{n}} \cdot \boldsymbol{\nabla\psi}$ – and sometimes they seem mysterious, but all that's really going on is that shadows get longer in the evening.[37] Note that the obliquity factors apply at both the emitter and detector, or alternatively in both the pupil and the image. A drinking straw's worth of flux requires $\sec(\theta_{\text{src}})$ times more area on the source, and is spread out over $\sec(\theta_{\text{det}})$ times more area on the detector.

1.3.16 Solid Angle

Waves expand as they propagate. However, as in the ray model, a given wave's angular spread is asymptotically constant as $R \to \infty$, or (equivalently) as we Fourier transform to $\mathbf{k}$ space or pass through a lens to the pupil. A plane angle is measured between two straight lines, so its measure doesn't depend on how far out you go. A cone has the same property in three dimensions, leading to a natural generalization: *solid angle*. This is shown using pupil coordinates (u, v) in Figure 1.10.

The measure of a plane angle is the arc length cut out by the angle on the unit circle, so we define the solid angle of a cone to be the area it cuts out of the unit sphere. (Note that the cone need not be circular in cross-section, or convex in outline,

35 Siegman, A.E. (1986). *Lasers*, University Science Books, 786–797.

36 Note that we're using field amplitudes and not intensity here. A similar trick is used in many-body quantum theory, where in a system in thermal equilibrium, a Boltzmann factor $\exp(-E/kT)$ is treated as an offset on the imaginary time axis.

37 At the risk of making things less clear again, when calculating the fields at some surface due to some incident light distribution, you have to multiply by $1/\cos\theta$ – see Section 9.3.5.

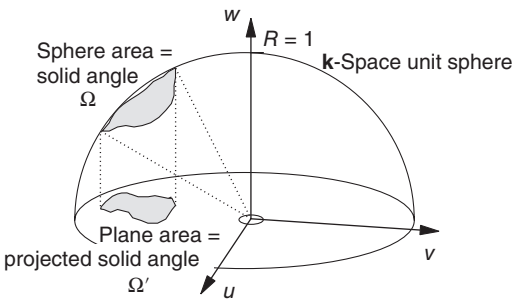


Figure 1.10 Definition of solid angle Ω and projected solid angle Ω′.

or even be a single glob – it just needs to have a shape that's independent of radial distance from its origin.) This area is of course

$$\Omega = \iint_{\text{cone}} \sin\theta \, d\theta \, d\phi, \tag{1.64}$$

where θ is the polar angle (measured from the surface normal) and ϕ is the azimuth (angle in the horizon plane). In optics, we're normally calculating the flux in or out of a surface, so we have to worry about obliquity. Obliquity is nothing deep or difficult to understand, as we just saw – when a beam of light hits a surface at an angle θ off normal, the illuminated patch is stretched out by sec θ, just as afternoon shadows of vertical objects lengthen as tan(θ). Mathematically, the outward flux through each bit of surface **dA** is $\mathbf{P} \cdot \mathbf{dA}$. It simplifies matters if we fold the obliquity into the quoted solid angle, so we usually work with the *projected solid angle* Ω', where

$$\Omega' = \iint_{\text{cone}} \sin\theta \cos\theta \; d\theta \; d\phi. \tag{1.65}$$

To crispen this idea up, consider a circular angular pattern of half-angle ψ around the surface normal, i.e. one that covers the angular disc $\theta < \psi$, $0 <= \phi < 2\pi$. Its solid angle is $\Omega = 2\pi(1 - \cos\psi) = \pi(\psi^2 - \psi^4/12 + \cdots)$ and its projected solid angle is $\Omega' = \pi\sin^2\psi = \pi(\psi^2 - \psi^4/3 + \cdots)$. Conveniently, if $n = 1$, then $\Omega' = \pi(\text{NA})^2$, which is a useful and easily remembered rule.

A Lambertian surface (one that has no preferred direction), emits into π steradians (the projected solid angle of a hemisphere). Optics folk tend to be loose about the distinction between Ω and Ω′, but it isn't hard to keep straight – if the emitter or receiver is a surface, there's obliquity to worry about, so use Ω′; if not (as in gas spectroscopy, for instance) use Ω. In the usual low-NA situations, the two are equivalent for practical purposes. There are two cautions to keep in mind: first, be careful if the surface isn't flat – it's the angle between the local surface normal and the light rays that matters. Second, both solid angle and obliquity are far-field concepts, so near a focus we have to use the plane-wave decomposition of the field to get the right answer.

1.3.17 Étendue: How Much Light Can I Get?

The first thing that an optical system has to be able to do is transmit light. Apart from solar telescopes, electro-optical systems are limited at least some of the time by how much light they can emit, collect, or detect. Figuring out how much you have and how much more you can get is the aim of radiometry. In Section 1.3.12, we saw that a given light beam can be focused into a smaller area, but only at the price of increasing its numerical aperture. Since $\sin\theta$ cannot exceed unity, a given beam cannot be focused arbitrarily tightly. Some beams can be focused better than others; for example a beam from an incandescent bulb cannot be focused as tightly as one from a laser. The difference is in their degree of spatial coherence.

The spatial coherence of a beam is a measure of how well its different components (Fourier or real-space) stay in phase with each other. This is revealed by how deep the interference fringes are when different components are made to interfere with one another, as in Young's slit experiment (there's more on this in Section 2.5). The theory of imaging with partially coherent light is discussed by Goodman, Born & Wolf, and others, and is beyond the scope of this book. As a practical matter, we usually want spatial coherence low enough to eliminate fringes in a full-field (i.e. not scanning) imaging system and high enough not to limit our ability to focus it on our area of interest. The *coherence area* of an optical field gives an idea of how far apart the slits can be and still have interference.

Conversely, one important attribute of an optical system is how well it can cope with low-coherence sources. To transmit the most light from such sources, the system needs both a large area and a large angular acceptance. The figure of merit for this attribute is called the *étendue* and is given by

$$E = n^2 A\Omega', \tag{1.66}$$

where A is the clear area and n is the refractive index of the medium in which the projected solid angle Ω' is measured. It's usually just written $A\Omega'$, which assumes that $n = 1$, but we'll carry the n along explicitly. For on-axis circular pupils (the usual case), $E = n^2 A\pi(\mathrm{NA})^2$. This quantity is invariant under magnification, which increases A while decreasing Ω' proportionately, and also under refraction.[38] Étendue is a purely geometric property, which explicitly neglects the transmittance of the optical system. This is fine as long as this is reasonably uniform up to the edges of A and Ω'. It is less useful with systems whose transmittance is a strong function of angle, e.g. those with high-index dielectric interfaces. The *weighted* étendue is not preserved on passing through a succession of such elements, so the transmittance must be expressed as a function of position and angle, and carried along mathematically. Étendue is related to the statistical mechanics notion of phase space volume, and the conservation of étendue is the optical analog of the conservation of phase space volume by adiabatic processes. (The Lagrange invariant is another way of stating the same thing, and the *ABCD* matrices of Section 1.3.12 are all unitary, which is the paraxial version.)

With a given low-coherence source, any two lossless optical systems with the same étendue will pass the same total optical power, if the source is matched to their characteristics with an appropriate magnification. Any mismatch will reduce the power actually transmitted. A corollary is that the étendue of a system stays the same if you send the light back through the other way. The étendue of an optical system cannot be larger than that of its poorest component and can easily be worse due to mismatch. This is worth keeping in mind, for example in the choice of polarizing prisms; types relying on TIR (such as the Glan–Taylor) have much smaller acceptance angles than those relying on double refraction (such as Wollastons), so a bigger prism can have a smaller étendue.

Example 1.8 ***Étendue and NA Mismatch***
Consider coupling sunlight into a 100×, 0.95 NA microscope objective ($f = 2$ mm, FOV radius = 50 μm). If we shine sunlight (9 mrad angular diameter) in the small end,[39] we get an effective $n^2 A\Omega'$ of $\pi(0.005\ \mathrm{cm})^2(\pi(4.5\ \mathrm{mrad})^2) = 5 \times 10^{-9}\ \mathrm{cm}^2\ \mathrm{sr}$. If we turn it around, the étendue is unaltered, but we get the 6 mm diameter back element instead. The angular acceptance on the exit pupil is a few degrees, which is still several times larger than the Sun's angular size; thus, we don't lose anything, so the effective $n^2 A\Omega'$ goes up by a factor of 3600 to $1.8 \times 10^{-5}\ \mathrm{cm}^2\ \mathrm{sr}$ – and the source is still mismatched.

1.3.18 What Is "Resolution"?

The classical definitions of Rayleigh and Sparrow specify that two point-like objects of equal brightness are resolved if their images are separated by a defined multiple of the diameters of their diffraction discs. This sort of definition is reasonably adequate for photographic detection, where the photon statistics do not significantly reduce the precision of the measurement.

With modern detectors, it is impossible to specify the resolution of an optical system unambiguously when signal-to-noise (SNR) considerations are absent. For example for a two-point object, one can model the image as the sum of two Airy patterns, whose locations and intensities are parameters. By fitting the model to the observed data, the positions and intensities of the two sources can be extracted. With a high enough SNR and a sufficiently accurate knowledge of the exact imaging characteristics of our systems, there is no clear limit to the two-point resolution of an optical system as defined in this way. Optical lithography is another example where the "resolution limit" has repeatedly turned out not to be where it was expected, largely on account of the very high contrast of photoresist and, recently, phase shift masks, computational mask design, and multiple exposures to exploit polarization and proximity effects.

What we really mean by *resolution* is the ability to look at an object and see what is there, in an unambiguous way that does not depend on our choice of model. This model-independent imaging property does degrade roughly in line with Rayleigh and Sparrow, but it is a much more complicated notion than simple two-point resolution. Most of the disagreement surrounding the subject of resolution is rooted here.

---

38 It's easiest to visualize this applying at images, where light from a given object point arrives at the corresponding image point from all angles. At out-of-focus planes, you have to calculate the range of angles illuminating each point, but the law still works.
39 That is, the side normally toward the sample.

1.4 Detection

To calculate what an instrument will detect, we need to know how to model the operation of a photodetector. Fortunately, this is relatively simple to do, providing that the detector is reasonably uniform across its sensitive area. From a physical point of view, all detectors convert optical energy into electrical energy, and do it in a *square-law* fashion – the electrical power is proportional to the square of the optical power, with a short time average through the detector's impulse response. (There's lots more on square law detection in Chapter 3, especially Section 3.2.1.)

Throughout the rest of this chapter, we will normalize the scalar field function ψ so that the (paraxial) power function $\psi\psi^*$ has units of watts per square meter.

A general square-law detector with an input beam $\psi(\mathbf{x})$ and a responsivity $\mathscr{R}$ will produce an output signal S given by

$$S(t) = \mathscr{R} \iint \langle \psi(\mathbf{x}, t)^* \hat{\boldsymbol{n}} \cdot \boldsymbol{\nabla}\psi(\mathbf{x})/ik \rangle \, d^2x \tag{1.67}$$

which for small NA is

$$S(t) = \mathscr{R} \iint \left\langle |\psi(\mathbf{x}), t)|^2 \right\rangle d^2x, \tag{1.68}$$

where angle brackets denote time averaging through the temporal response of the detector and the integral is over the active surface of the detector. The gradient $\nabla\psi$ is parallel to the local direction of propagation (see Section 9.2.3), and the dot product supplies the *obliquity factor* as we saw in Section 1.3.16. If the detector is seriously nonuniform, the responsivity becomes a function of $\mathbf{x}$, so $\mathscr{R}(\mathbf{x})$ must be put under the integral sign.[40]

The square-law behavior of detectors has many powerful consequences. The first is that all phase information is lost; if we want to see phase variations, we must convert them to amplitude variations before the light is detected. Furthermore, provided that no light is lost in the intervening optical components (watching vignetting especially), the detector can in principle be placed anywhere in the receiving part of the optical system because the time averaged power will be the same at all positions by conservation of energy. This has great practical importance because we may need to use a small detector in one situation, to minimize dark current, ambient light sensitivity, or cost, and a large one in another to prevent saturation and attendant nonlinearity due to high peak power levels. The small detector can be put near focus and the large one far away. This freedom applies mathematically as well; provided once again that no additional vignetting occurs, (1.68) can be applied at an image, a pupil, or anywhere convenient.[41] This becomes very useful in interferometers.

As in all interactions of light with matter, the surface properties of the detector and their variation with position, polarization, and angle of incidence are important. Fortunately, detector manufacturers endeavor to make their products as easy to use as possible so that the worst nonuniformities are eliminated, and in addition, by the time the light gets to the detector, its numerical aperture is usually reduced sufficiently that obliquity factors and dependence on overall polarization are not too serious. As usual, they can be put in by hand if needed, so we'll continue to use the scalar model and neglect these other effects.

There are a fair number of head-scratchers associated with square-law detection. We'll talk more about it in Section 3.2.

1.5 Coherent Detection

1.5.1 Interference

An interferometer is nothing more than a device that overlaps two beams on one detector, *coherently*, rather than combining the resulting photocurrents afterward, *incoherently*. Coherent addition allows optical phase shifts between the beams to give rise to signal changes. In many applications, the two beams are different in strength and the weaker one carries the signal information. Consequently, they are often referred to as the signal and local oscillator (*LO*) beams, by analogy with superheterodyne radios. Coherent detection gives the optical fields the chance to add and subtract before the square-law is applied so that the resulting photocurrent is

$$\begin{aligned} i(t) &= \mathscr{R} \iint_{\text{det}} \left| \psi_{\text{LO}}(\mathbf{x}) e^{-i(\omega_{\text{LO}} t + \phi_{\text{LO}}(\mathbf{x},t))} + \psi_S(\mathbf{x}) e^{-i(\omega_S t + \phi_S(\mathbf{x},t))} \right|^2 dA \\ &= i_{\text{LO}} + i_S + i_{\text{AC}} \end{aligned} \tag{1.69}$$

40 If your NA or obliquity is unusually high at the detector, you may also need to consider polarization.

41 This is exact and not a Debye approximation.

assuming that the beams are polarized identically (if there are components in the orthogonal polarization state, they add in intensity, or equivalently in i). The individual terms are

$$i_{\mathrm{LO}} = \mathscr{R} \iint_{\mathrm{det}} d^2x \psi_{\mathrm{LO}} \psi_{\mathrm{LO}} *, \tag{1.70}$$

$$i_S = \mathscr{R} \iint_{\mathrm{det}} d^2x \psi_S \psi_S *, \text{and} \tag{1.71}$$

$$\begin{aligned} i_{\mathrm{AC}} &= 2\mathscr{R}\,\mathrm{Re}\left\{ \iint_{\mathrm{det}} d^2x \psi_{\mathrm{LO}} \psi_S * \right\} \\ &= 2\mathscr{R}\,\mathrm{Re}\left\{ \exp(-i\Delta\omega t) \iint_{\mathrm{det}} \left|\psi_{\mathrm{LO}}(\mathbf{x})\right| \left|\psi_s(\mathbf{x})\right| \exp(i\Delta\phi(\mathbf{x},t)) dA \right\}. \end{aligned} \tag{1.72}$$

The first two terms, i_{LO} and i_S, are the photocurrents the two beams would generate if each were alone. The remaining portion is the *interference term*. It contains information about the relative phases of the optical beams as a function of position. The interference term can be positive or negative, and if the two beams are at different optical frequencies, it will be an AC disturbance at their difference frequency $\Delta\omega$. If the two beams are superposed exactly and have the same shape (i.e. the same relative intensity distributions, $\mathbf{k}$ vector, focus, and aberrations), ψ_{LO} and ψ_S differ only by a constant factor. This can be taken out of the integral, so the interference term becomes

$$i_{\mathrm{AC}} = 2\sqrt{i_{\mathrm{LO}}\, i_S} \cos(\Delta\omega t + \phi). \tag{1.73}$$

This is remarkably powerful, as we'll see.

Aside: Fringe Visibility

Looking at the light intensity on the detector (or on a sheet of paper), we can see a pattern of light and dark fringes if the light is sufficiently coherent. These fringes are not necessarily nice looking. For laser beams of equal strength, they will go from twice the average intensity to zero; for less-coherent sources, the fringes will be a fainter modulation on a constant background. The contrast of the fringes is expressed by their visibility V,

$$V = \frac{I_{\mathrm{max}} - I_{\mathrm{min}}}{I_{\mathrm{max}} + I_{\mathrm{min}}} \tag{1.74}$$

which we'll come back to in Section 2.5 in the context of coherence theory.

1.5.2 Coherent Detection and Shot Noise: The Rule of One

Application of coherent detection to improve the SNR is covered in Section 3.10.7. There are three key observations to be made here: coherent detection is extremely selective, preserves phase information, and provides noiseless signal amplification.[42] These three properties give it its power. If the two beams are exactly in phase across the entire detector, the amplitude of the interference term is twice the square root of the product of the two DC terms

$$i_{\mathrm{AC}}(\mathrm{peak}) = 2\sqrt{i_{\mathrm{LO}} i_S}. \tag{1.75}$$

If i_S is much weaker than i_{LO}, this effectively represents a large amplification of ψ_S. The amplification is completely noiseless – the LO shot noise is

$$i_{N\mathrm{shot}} = \sqrt{2e i_{\mathrm{LO}}} \tag{1.76}$$

which is exactly the root mean square (RMS) value of i_{AC} when i_S is 1 electron per second (the noise current is down by $\sqrt{2}$ due to the ensemble average over 2π phase). Thus, with a quantum efficiency $\eta = 1$, a signal beam of 1 photon/s is detectable at 1σ confidence in one second in a 1 Hz bandwidth, which is a remarkable result – in other words, bright-field measurements can be made to the same sensitivity as dark-field measurements.[43]

This leads us to formulate the Shot Noise Rule of One: *One* coherently added photon per *One* second gives an AC measurement with *One* sigma confidence in a *One* hertz bandwidth. (See Sections 1.8.1 and 13.1 for more on AC vs. DC measurements.)

---

42 Jelalian, A.V. (1992). *Laser Radar Systems*, 33–41. Boston: Artech House.

43 There is one fine point: in dark field there is no shot noise when there is no signal, whereas here the uncertainty is equally distributed between the $|0>$ and $|1>$ states. The average variance is the same in the two cases, just differently distributed.

Aside: Photons Considered Harmful
Thinking in terms of photons is useful in noise calculations but pernicious almost everywhere else – see Section 3.2.2.

1.5.3 Spatial Selectivity of Coherent Detection

If the phase relationship is not constant across the detector, fringes will form, so the product $E_{\mathrm{LO}}E_s^*$ will have positive and negative regions; this will reduce the magnitude of the interference term. As the phase errors increase, the interference term will be reduced more and more, until ultimately it averages out to nearly zero. This means that a coherent detector exhibits gain only for signal beams that are closely matched to the LO beam, giving the effect of a matched spatial filter plus a noiseless amplifier.

Another way to look at this effect is to notionally put the detector at the Fourier transform plane, where the two initially uniform beams are transformed into focused spots. A phase error that grows linearly across the beam (equally spaced fringes) corresponds to an angular error, which in the transform plane means that the two focused spots are not concentric. As the phase slope increases, the spots move apart so that their overlap is greatly reduced. Ultimately, they are entirely separate and the interference term drops to zero. Mathematically, these two are equivalent, but physically, they generally are not.

If the detector is placed at a focused spot, the local photocurrent density can be so large as to lead to pronounced nonlinearity; this is much less of a problem when the beams fill a substantial fraction of the detector area. On the other hand, two spots that do not overlap will not give rise to any interference term whatsoever, which is not in general true of two broad beams exhibiting lots of interference fringes; even if the broad beams are mathematically orthogonal, small variations in sensitivity across the detector (e.g. due to dust, etalon fringes, or nonuniformity across the device itself) will prevent their interference pattern from averaging to exactly zero.

It is hard to say exactly how serious this effect is in a given case, as it depends strongly on the details of the sensitivity variations. Sharp, strong variations (e.g. vignetting) will give rise to the largest effects, while a smooth center-to-edge variation may do nothing at all noticeable. If the application requires >40 dB (electrical) selectivity between beams at different angles, consider changing focus to separate them laterally, or relying on baffles or spatial filters as well as fringe averaging.

1.5.4 Optical Modes, Antennas, and Thermodynamics

Coherent detection is very much analogous to a radio antenna, one of the very many fruitful analogies between electro-optics and early radio that we'll talk about in Section 13.11.[44] An antenna with one pair of feed terminals can interrogate only one mode of the electromagnetic (EM) field, in one polarization state, per frequency. A single polarized mode is sort of the quantum of étendue:

$$n^2 A\Omega = \lambda^2/2. \tag{1.77}$$

It's just like single-mode fiber, except that SMF accepts any polarization, so you get another factor of 2, i.e. its étendue is λ^2. (You need two sets of terminals to do that with an antenna.) For instance, ordinary SM-28 fiber at 1.55 μm has a mode diameter of 9 μm and an NA of 0.11 in air, so for a check,

$$n^2 A\Omega/\lambda^2 = [\pi(4.5\ \mu\mathrm{m})^2][\pi 0.11^2]/(1.55\ \mu\mathrm{m})^2 = 1.007 \tag{1.78}$$

(quoted to more significant figures than we're entitled to). The second law of thermodynamics, the one that says heat doesn't spontaneously flow from cold to hot, can be formulated in terms of modes (see Section 2.4.2).

Optical modes are useful when thinking about coherent detection because what the coherent detector generates is precisely the overlap integral between ψ_{LO} and ψ_S. Distinct modes are orthogonal, i.e. the overlap integral between them is zero. Thus, each mode interferes only with itself, and the étendue limit (1.77) defines the tradeoff of area, angular acceptance, and depth of focus. (Segmenting the detector can sometimes help.) That λ^2 is the main reason that even coherent laser radar is so much less sensitive than the radio version: for the same geometry, the intercepted area goes down by factors[45] like 10^{12}.

---

44 See, e.g. Hobbs, P.C.D. et al. (2007). Efficient waveguide-integrated tunnel junction detectors at 1.6 μm. *Opt. Express* 15 (25): 16376–16389. http://www.opticsinfobase.org/oe/viewmedia.cfm?uri=oe-15-25-16376&seq=0.

45 Incoherent detection is worse because the signal power goes down quadratically with decreasing optical power. Increased photon energy doesn't help either: fewer photons per watt means that the shot noise of the background becomes obnoxious.

1.6 Interferometers

1.6.1 Two-Beam Interferometers

Two-beam interferometers implement the scheme of Section 1.5 in the simplest way: by splitting the beam into two with a partially reflecting mirror, running the two through different paths, and recombining them. Figure 1.11 shows the heavy lifters of the interferometer world, the Michelson and Mach–Zehnder. Mach–Zehnders are more common in technological applications, because the light goes in only one direction in each arm, so it's easier to prevent back-reflections into the laser. On the other hand, a Michelson is the right choice when robust alignment is needed because one or both mirrors can be replaced by corner cubes (don't tell anyone, but if the cubes are offset from the beam axis, which is usually a very good idea, that's really a skinny Mach–Zehnder). Example 1.12 shows an intensive use of a corner cube type interferometer. Michelsons are a bit easier to align because sending the beam back on itself is an easy criterion to use.

An interferometer is intrinsically a four-port device; light is steered between the output ports by interference. If the two beams are perfectly coherent with one another and have the same polarization, the output powers P_{O+} and P_{O-} from the two output ports are

$$P_{O\pm} = P_1 + P_2 \pm 2\sqrt{P_1 P_2}\cos\phi, \tag{1.79}$$

where P_1 and P_2 are the split-beam powers and ϕ is the phase angle between them. The sign difference comes from beam 1 being reflected and beam 2 transmitted going into port + and vice versa for port −.

Aside: Interferometers and Polarization

It's usually best to use polarizing beam splitters in interferometers because that saves a lot of light (see Section 4.8.1). However, that brings up the issue of polarization. A polarizing beam splitter reflects *s* polarization and transmits *p*. In a Michelson, for instance, the light that gets reflected on the transmit path gets reflected on the receive path, and similarly the transmitted light gets transmitted again, so both wind up going back to the laser, which is usually not what you want. In a Mach–Zehnder, that doesn't send the beams back to the laser, but they're still orthogonally polarized and so need a bit of help to make them interfere. This can be a bit puzzling if you don't follow the polarization around the optical path. See Example 1.12 for ways to handle this.

1.6.2 Multiple-Beam Interferometers: Fabry–Perots

If instead of splitting the light into multiple paths, we just take two partially reflecting mirrors and put them next to each other, parallel, we get a *Fabry–Perot* interferometer. The multiple reflections give Fabry–Perots much greater selectivity for a given size, at the expense of far greater vulnerability to mirror errors and absorption. We'll go through the math in Section 5.4, but the upshot is that a lossless plane-mirror F-P whose mirrors have reflectance R and are spaced d apart in a medium of index n has a total transmission:

$$T_{\mathrm{FP}} = \frac{1}{1 + \frac{4R}{(1-R)^2}\sin^2\left(nk_0 d\cos\theta\right)}, \tag{1.80}$$

where θ is the angle of incidence of the beam on the mirrors (inside the medium). With high-reflectance mirrors, $1 - R$ is small, so T_{FP} consists of spikes at the nulls of the sine function, i.e. multiples of $\Delta\nu = 1/(2n\,d)$, the *free spectral range*

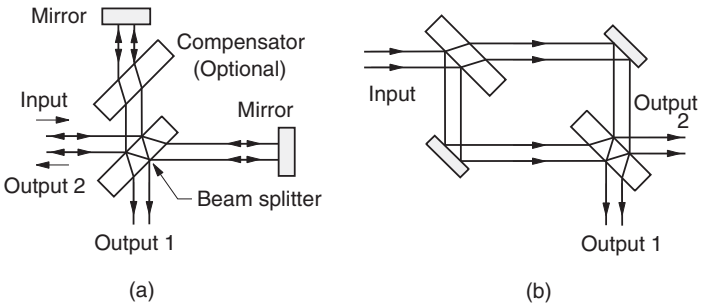


Figure 1.11 Workhorse two-beam interferometers: (a) Michelson; (b) Mach–Zehnder. The compensator plate in (a) more or less eliminates the effects of dispersion in the glass of the beam splitter, which will smear out the white-light fringes otherwise, and also reduces the effect of finite aperture.

(FSR). The full width at half maximum (FWHM) of the peaks is FSR/$\mathscr{F}$, where $\mathscr{F}$ is the *finesse*. Although $\mathscr{F}$ is nominally $(\pi\sqrt{R})/(1-R)$, it is really a measured quantity because mirror flatness errors are usually the limiting factor. If the mirrors have RMS error δ in waves, that limits the finesse to

$$\mathscr{F}_{\max} < 1/(2\delta), \tag{1.81}$$

which is a pretty serious limitation most of the time – achieving a finesse of 100 requires mirror accuracy and alignment precision of better than $\lambda/200$. Real F-Ps have a peak T less than 1 (sometimes a lot less), and their total reflectance R_{FP} is less than $1 - T_{\mathrm{FP}}$. Recently, fabrication precision and coating quality have advanced to the point where finesse values of 2×10^5 or higher can be obtained, at a price.

Inside a F-P, the field is enhanced a great deal; the easiest way to calculate it is to notice that the forward and backward propagating fields inside the cavity are nearly equal, and that the transmitted power has to be T times the forward power. In a perfect F-P, that means that if P_{inc} is coming in, the forward power inside is $T_{\mathrm{FP}}/(1-R) \cdot P_{\mathrm{inc}}$.

Aside: Other Multiple-Beam Interferometers
The other classes of multiple beam or "amplitude splitting" interferometers are Fizeau (linear fringes from a wedge-shaped gap), Newton (circular fringes from a sphere touching a flat), and Haidinger (Fabry–Perot with an extended source such as a mercury lamp) are all basically variations on the Fabry–Perot system, so it probably isn't worth memorizing them all.

1.6.3 Focused-Beam Resonators

Fabry–Perots can have variable or fixed spacing; a fixed F-P is called an *etalon*. Etalons can be tuned over a small range by tipping them, but the finesse drops pretty fast if you go too far since the Nth reflections start missing each other completely.

As the finesse goes up, even small amounts of diffraction become an increasing difficulty. Thus, the highest-finesse Fabry–Perots use at least one curved mirror. Like laser resonators, they must match a certain Gaussian beam mode, which requires a particular combination of mirror curvature and beam diameter for a given λ and cavity length. For best performance, you have to match the incoming wave to the spherical wave cavity mode. This is a huge pain if you have to do it manually more than once. Fortunately, single-mode fiber coupled ones are available – buy that kind if you possibly can. Otherwise, not only do you face critical matching problems, but the minor pointing instability of the laser will turn into noise that cannot easily be removed. Fibers have their problems, but very high finesse focused F-Ps are much, much worse.

Aside: Confocal Cavities and Instability
It might seem that the ideal optical resonator would be a confocal cavity, where the two mirrors' centers of curvature coincide at the center of the cavity. This is OK for spectrum analyzers but not for lasers. The reason is quite interesting. Such a cavity is a portion of a single sphere, and there is no special direction in a sphere – any direction is as good as any other, hence a confocal resonator has no stable axis. A tilt of ε radians in one mirror produces a shift of the resonator axis of $\delta\theta \approx \varepsilon(L/(\Delta L))$, where L is the distance between the mirror vertices and ΔL is the distance between their foci – which goes to ∞ as $\Delta L \to 0$ so that the laser pointing stability goes right in the tank. The NA of the resonant mode depends on how far off confocal the cavity is – resonant wavefronts will coincide with the cavity mirror surfaces, so NA = 0 for planar mirrors, NA = 1 for confocal mirrors, and in between, the NA can be backed out from the equation for $R(z)$ in Table 1.1. (Should ΔL be positive or negative?)

1.7 Photon Budgets and Operating Specifications

1.7.1 Basis

Photons are like dollars: a certain number are needed for the job at hand, and they're easier to lose than to get back. Thus, the idea of a budget applies to photons as to finances, but it is more complicated in that not all photons are useful – as though we had to budget a mixture of green and purple dollars. A photon budget is an accounting of where photons come from, where they go, and how many are expected to be left by the time they are converted to electrons by the detector.[46]

46 "Photon budget" is really shorthand for an overall feasibility calculation based on first principles and various rules of thumb, but SNR, selectivity and stability are the most important aspects.

Also like the other kind, people sometimes don't even expect to be on budget; they settle for "this was the best we could do…I'm not sure what was the problem." In the author's experience, it is possible to achieve a SNR within 3 dB of budget almost always, and 1 dB most of the time. Don't give up; this theory stuff really works.

On the other hand, the budget must be realistic too. Don't try to measure anything in a bandwidth of less than 10 Hz, unless you have lots of cheap graduate students, and remember that you need a decent SNR to actually do anything. Sensitivity limits are frequently given as *noise equivalent power* (NEP) or *noise equivalent temperature difference* (NEΔT or NETD), meaning the amount of signal you need in order to reach a SNR of 1 (0 dB), or equivalently a confidence level of 1σ (68%). Don't let this convince you that a SNR of 1 is useful for anything because it isn't. Generally, for any reasonable measurement you need a SNR of at least 20 dB – even to do a single go/no-go test with an acceptable false call probability, you'll need at least 10 or 15 dB ($3–5\sigma$). Tolerable images need at least 20 dB SNR; good ones, about 40 dB. Just how large a SNR your measurement has to have is a matter of deepest concern, so it should be one of the first things on the list. Don't rely on a rule of thumb you don't understand fully, including this one. There's lots more on this topic in Section 13.6.

Arriving at a photon budget and a set of operational specifications is an iterative process, as the two are inseparably linked. As the concepts are simple, the subtlety dwells in the actual execution; we will therefore work some real-life examples. The first one will be a shot-noise limited bright field system; the second, a background limited dark field system, and the third, an astronomical CCD camera. Before we start, you'll need to know how to combine noise sources (see Section 13.6.10) and to think in decibels.

Aside: Decibels

One skill every designer needs is effortless facility with decibels. There are just two things to remember: first, *decibels always measure power ratios, never voltage*. $G(\text{dB}) = 10\,\log_{10}(P_2/P_1)$ (we'll drop the subscript "10" from now on and use "ln" for natural log). That formula with a "20" in it is a convenience which applies only when the impedances are identical, e.g. a change in the noise level at a single test point. This is because it assumes that the power goes like the voltage squared, which is true only if the impedance is constant. (If you don't remember this, you'll start thinking that a step-up transformer has gain.) Second, you can do quick mental conversions by remembering that a factor of 2 is 3 dB (since $\log(2) \approx 0.3010$), a factor of 10 is 10 dB, and ± 1 dB is $1.25\times$ or $0.8\times$ (since $10^{0.1} \approx 1.259$). For example, if you're looking at the output of an amplifier, and the level changes from 10 to 77 mV, that's about an 18 dB change: you get 80 from 10 by multiplying by 10 and dividing by 1.25 ($20 - 2$ dB, remembering that dB measure power), or by multiplying by two three times ($6 + 6 + 6 = 18$). Absolute power levels are usually specified in dBm, i.e. decibels with respect to 1 mW so that 1 μW is −30 dBm and 1 W is +30 dBm.[47]

Another useful property is that formulas quoted in decibels display the scaling laws in a nice explicit form. For instance, a resistor in a matched circuit produces a noise power of kT per hertz.[48] In the signal-processing world, this is usually quoted as (assuming $T = 300$ K)

$$P_{\text{noise}}(\text{dBm}) = -173.8 + 10\ \log \text{BW}, \tag{1.82}$$

which makes the bandwidth scaling obvious. One bad side effect of this is that a lot of the physics gets hidden in apparently mysterious fudge factors like that −173.8 dBm. We'll see some more of that in Section 14.11.8 in connection with A/D converter noise.

Example 1.9 ***Photon Budget for a Dual-Beam Absorption Measurement***

One way of compensating for variations in laser output is to use two beams, sending one through the sample to a detector and the other one directly to a second detector for comparison. A tunable laser, e.g. Ti:sapphire or diode, provides the light. The desired output is the ratio of the instantaneous intensities of the two beams, uncontaminated by laser noise, drift, and artifacts due to etalon fringes or atmospheric absorption. For small absorptions, the beams can be adjusted to the same strength, and the ratio approximated by their difference divided by a calibration value of average intensity,

$$\varepsilon = \frac{N_{\text{sig}}(\lambda, t)}{N_{\text{comp}}(\lambda, t)} - 1 \approx \frac{N_{\text{sig}}(\lambda, t) - N_{\text{comp}}(\lambda, t)}{\left\langle N_{\text{comp}}(\lambda) \right\rangle}, \tag{1.83}$$

where N_{sig} and N_{comp} are the photon flux (s^{-1}) in the beams. The total electrical power in the signal part is

$$P_{\text{sig}} = \left[\eta e (N_{\text{sig}} - N_{\text{comp}})\right]^2 R_L, \tag{1.84}$$

where η is the quantum efficiency.

47 In a 50-Ω system, 0 dBm is 632 mV p-p.

48 See Section 13.2.5 for why it isn't $kT/2$ per hertz.

In the absence of other noise sources, shot noise will set the noise floor:

$$i_{N\text{shot}} = e\sqrt{2\eta(N_{\text{comp}} + N_{\text{sig}})}. \tag{1.85}$$

For 1 mW per beam at 800 nm and $\eta = 1$, this amounts to a dynamic range (largest electrical signal power/noise electrical power) of 150 dB in 1 Hz, or a 1-σ absorption of three parts in 10^8. Real spectrometers based on simple subtraction are not this good, due primarily to laser noise and etalon fringes. Laser noise comes in two flavors, intensity and frequency; it's treated in Section 2.13.

Laser noise cancellers use subtraction to eliminate intensity noise and actually reach this shot noise measurement limit (see Sections 10.9.6 and 18.7.3). When frequency noise is a problem (e.g. in high-resolution spectroscopy), we have to stabilize the laser. Frequency noise also couples with the variations in the instrument's T vs. λ to produce differential intensity noise which in general cannot be canceled well. If the instrument's optical transmittance is T and the laser has a (one-sided) frequency modulation (FM) power spectrum $S(f_m)$ (which we assume is confined to frequencies small compared to the scale of T's variation), it will contribute RMS current noise i_{Nfa}

$$i_{Nfa}(f) = N\eta e S(f)\frac{dT}{d\nu}. \tag{1.86}$$

If this noise source is made small, and the absorption can be externally modulated, e.g. by making the sample a molecular beam and chopping it mechanically, the shot noise sensitivity limit can be reached fairly routinely. Note that this is not the same as a measurement accuracy of this order; any variations in T or drifts in calibration will result in a multiplicative error which, while it goes to zero at zero signal, usually dominates when signals are strong.

Example 1.10 ***Photon Budget for a Dark Field Light Scattering System***

Many systems, e.g. laser Doppler anemometers, detect small amounts of light scattered from objects near the waist of a beam. Consider a Gaussian beam with $P = 0.5$ mW and 0.002 NA at 633 nm. From Table 1.1, the 3-dB beam radius at the waist is 70 μm, and the central intensity is $2P\pi(\text{NA})^2/\lambda^2 = 3.1 \times 10^4$ W/m^2, which is 1.0×10^{23} photons/m^2/s. A sufficiently small particle behaves as a dipole scatterer, so that the scattered light intensity is zero along the polarization axis and goes as the sine squared of the polar angle θ. It will thus scatter light into 2π steradians, so if the total scattering cross section of the particle is σ_{tot}, the averaged scattered flux through a solid angle Ω (placed near the maximum) will be $F = 1.0 \times 10^{23}\sigma_{\text{tot}}\Omega/(2\pi)$. (Larger particles exhibit more angular structure in their scattering behavior, with a very pronounced lobe near the forward direction.) A 1 cm diameter circular detector at a distance of 5 cm from the beam waist will subtend a projected solid angle of $\Omega' = \pi(\text{NA})^2 = \pi(0.5/5)^2 \approx 0.03$. A particle crossing the beam waist will thus result in $N = 5 \times 10^{21}\sigma_{\text{tot}}$ photons/s.

If the detector is an AR-coated photodiode ($\eta \approx 0.9$) with a load resistor $R_L = 10$ MΩ, then in a time t (bandwidth $1/(2t)$ Hz), the 1-Hz[49] RMS Johnson noise current $(4kTB/R)^{1/2}$ is $(2kT/10^7)^{1/2}$, which is 0.029 pA or 1.8×10^5 electrons/s (which isn't too good).

From Section 13.6.19, we know that in order to achieve a false count rate R of 1 in every 10^6 measurement times, we must set our threshold at approximately 5.1 times the RMS amplitude of the additive Gaussian noise (assuming that no other noise source is present). Thus, a particle should be detectable in a time t if σ_{tot} exceeds

$$\sigma_{\text{min}} = \frac{1.8 \times 10^5}{5 \times 10^{21}}\frac{(5.1)}{\eta\sqrt{t}} = \frac{1.8 \times 10^{-16}}{\eta\sqrt{t}}, \tag{1.87}$$

where η is the quantum efficiency. A particle moving at 1 m/s will cross the 140 μm 3-dB diameter of this beam in 140 μs, so to be detectable it will need a scattering cross section of at least 1.6×10^{-14} m^2, corresponding to a polystyrene latex (PSL) sphere ($n = 1.51$) of about 0.2 μm diameter.

Now let's sanity-check our assumptions. For the circuit to respond this fast, the time constant $C_{\text{det}}R_L$ must be shorter than 100 μs or so, which limits the detector capacitance to 10 pF, an unattainable value for such a large detector. Even fast detectors are generally limited to capacitances of 50 pF/cm^2, so assuming the detector capacitance is actually 100 pF, we cannot measure pulses faster than 1–2 ms with our system as assumed. There are circuit tricks that will help considerably (by as much as 3600×, see Section 18.5.2), but for now, we'll work within this limit. If we can accept this speed limitation, and the accompanying ∼ 10× decrease in volumetric sampling rate, we can trade it in for increased sensitivity; a particle whose transit time is 2 ms can be detected at $\sigma_{\text{min}} = 4 \times 10^{-15}$ m^2. If the particle is much slower than this, its signal will start

49 *In* 1 Hz, not *at* 1 Hz.

to get down into the $1/f$ noise region, and non-Gaussian noise such as popcorn bursts will start to be important. Already it is in danger from room lights, hum, and so on.

If the detector is a photon-counting photomultiplier tube (PMT, whose quantum efficiency $\eta \approx 0.2$), the noise is dominated by the counting statistics, and the technical speed limitation is removed (see Section 3.5). However, PMTs have a certain rate of spurious *dark counts*. Dark counts generally obey Poisson statistics, so if we assume a mean rate $N_{\text{dark}} = 200$ Hz, then in a 140 μs measurement, the probability of at least one dark count is $200 \times 140\ \mu\text{s} \approx 0.028$. We are clearly in a different operating regime here, where the fluctuations in the dark count are not well described by additive Gaussian noise as in the previous case. From Section 13.6.20, the probability of a Poisson process of mean rate λ per second producing exactly M counts in t seconds is

$$P(M) = \frac{(\lambda t)^M e^{-\lambda t}}{M!}. \tag{1.88}$$

If we get 200 dark counts per second, then the probability that another will arrive within 140 μs of the first one is $(0.028)e^{-0.028} \approx 0.027$, so we expect it to happen $200 \times 0.027 \approx 5.5$ times a second. Two or more additional counts will arrive within 140 μs of the first one about 0.076 times per second, and three or more additional counts only 0.0007 times per second, so if we require at least four counts for a valid event, the false count rate will be one every 20 minutes or so, which is usually acceptable. The limiting value σ_{min} is then

$$\sigma_{\text{min}} \approx \frac{4}{5 \times 10^{21} \eta t} \approx \frac{8 \times 10^{-22}}{\eta t}. \tag{1.89}$$

which is about 3×10^{-17} m^2, nearly three orders of magnitude better than the photodiode, and sufficient to detect a PSL sphere of 0.08 μm (alas for particle counters, the signal goes as a^6). We can use the same Poisson statistics to predict the probability of detection, i.e. $P(\geq 4$ photons) as a function of σ_{tot}.

Besides detector capacitance, such measurements are often limited by the shot noise of light from the background, from imperfect beam dumps, or from molecular Rayleigh scatter (as in the blue sky), so that our super-simple photon budget is not a sufficient theoretical basis for the planned measurement. More detail is available in Chapter 10. In addition, we have here assumed a very simple deterministic model for signal detection in noise; any event whose nominal detected signal is greater than our threshold is assumed to be detected. This assumption is unreliable for signals near the threshold and is dealt with in a bit more detail in Section 13.6.19. Finally, we assumed that our noise was limited by the use of a boxcar averaging function of width t, equal to the 3 dB width of the pulse. Even with *a priori* knowledge of the arrival time of the particle, this is not the optimal case. While the optimum is quite broad, still if the particles are moving much faster or slower than we anticipate, a fixed averaging window may be far from optimal. This is the topic of Section 13.8.10.

Example 1.11 ***Photon Budget for an Astronomical CCD Camera***

An astronomical telescope gathers photons from sources whose brightness cannot be controlled. It also gathers photons from Earth, due to aurora, meteors, cosmic ray bursts, scattered moonlight, and human artifacts such as street lights. Extraneous celestial sources such as the zodiacal light and the Milky Way are also important in some parts of the sky. It represents an interesting signal detection problem: the signals that it tries to pull out of random noise are themselves random noise. The inherent noisiness of the signal is of less consequence in the optical and infrared than in the radio region. The brightness of sky objects is quoted in logarithmic relative *magnitudes* , with respect to standard spectral filters. A look in Allen's *Astrophysical Quantities* (affectionately known as "AQ") reveals that in the "visible" (V) filter band, centered at 540 nm in the green, a very bright star such as Arcturus or α Centauri has a magnitude $m_V = 0$, and that such an object produces a total flux at the top of the atmosphere of about 3.8 nW/m^2, of which about 80% makes it to the Earth's surface. A first magnitude star ($m_V = 1.0$) is 100 times brighter than a sixth magnitude star, which is about the limit of naked-eye detection in a dark location. A one-magnitude interval thus corresponds to a factor of $100^{0.2} \approx 2.512$ in brightness.[50]

Even in a dark location, the night sky is not perfectly dark; its surface brightness in the V band is about 400 nW/m^2/sr, which at 2.3 eV/photon is about 1×10^{10} photons/s/m^2/sr. An extended object such as a galaxy will be imaged over several resolution elements of the detector, whereas (unless we have enough pixels to resolve the point-spread function of the telescope) a point-like object such as a star will usually be imaged onto a single detector element. Without adaptive optics, the turbulence of the atmosphere during the (long) measurement time limits the resolution to the size of a *seeing disc* of

50 The magnitude scale goes back to the ancient Greeks – the numerical factor is weird because the log scale was bolted on afterward and tweaked to preserve the classical magnitudes while being reasonably memorable.

diameter 0.25″ (arc seconds) on the best nights, at the best locations, with 1″ being more typical, and 3″–5″ being not uncommon in backyard observatories. Thus, with enough pixels even a stellar object produces an extended image, and the SNR will be limited by the fluctuations in the sky brightness in a single pixel. Each pixel subtends the same solid angle Ω on the sky, so the mean photoelectron generation rate per pixel is

$$n_{\text{tot}} = \eta\Omega A \left(J_{\text{sky}} + J_{\text{star}}\right), \tag{1.90}$$

where A is the area of the telescope objective, N is in electrons, n is in electrons/s, and J_{star} and J_{sky} are the photon angular flux densities in photons/s/m^2/sr, respectively, from the star and from one pixel's worth of sky background.

There are two classes of noise source in the measurement: counting statistics of the sky photoelectrons and of the dark current, which go as $(nt)^{1/2}$, and the (fixed) readout noise q_{RO}. The electrical SNR is thus

$$\text{SNR} = \frac{(J_{\text{star}}\ t\ \eta\Omega A)^2}{\Delta n_{\text{RO}}^2 + n_{\text{dark}}t + t\eta\Omega A\left(J_{\text{star}} + J_{\text{sky}}\right)}. \tag{1.91}$$

CCD pixels have a fixed *full well* capacity B, ranging from about 5×10^4 electrons for the ones used in a handheld camera to 10^6 for a scientific device. Thus, the maximum exposure time for a perfect device is equal to eB/i_{sky}, which is on the order of a week so that the well capacity is not a limitation for dim objects. The noise-equivalent photon flux density (NEΔJ) is the number of photons per second required to achieve a SNR of 0 dB in the measurement time. Setting the SNR in (1.91) to 1 yields a quadratic equation for J_{star}, whose solution is

$$\text{NE}\Delta\text{J}_{\text{star}} \approx \frac{1}{\eta\Omega A}\sqrt{\frac{n_{\text{dark}} + \eta\Omega A J_{\text{sky}}}{t} + \frac{\Delta n_{\text{RO}}^2}{t^2}}. \tag{1.92}$$

A 100 cm telescope with a 40% central obstruction (due to the secondary mirror) has an area $A = 0.66$ m^2, and an angular resolution of about 0.14 arc seconds. A commercially available V filter has a transmittance of about 0.72 at the nominal peak.[51] Assuming that the telescope itself has an efficiency of 0.8, due primarily to absorption in uncoated aluminum primary and secondary mirrors, and is used with a cooled, back-surface illuminated CCD with quantum efficiency $\eta = 0.85$ and $B = 1.8 \times 10^5$, the end-to-end quantum efficiency of the telescope system is $\eta \approx 0.5$. A good cooled CCD has a dark current of around one electron per second per pixel, and RMS readout noise of about five electrons. With this apparatus,

$$\text{NE}\Delta\text{J} \approx \frac{3}{\Omega}\sqrt{\frac{1 + 3.3 \cdot 10^{11}\Omega}{t} + \frac{5^2}{t^2}}. \tag{1.93}$$

This is dominated by readout noise at short exposures. At longer times, dark current fluctuations dominate for very small values of Ω but sky background fluctuations for larger Ω. The largest SNR improvement comes from increasing integration time, as is shown in Figures 1.12–1.15, which are pictures of the galaxy M100 taken by Dennis di Cicco using an 11-inch Schmidt–Cassegrain telescope at $f/6.2$ and an SBIG ST-6 CCD camera.[52]

With an exposure time of an hour, and a 3×3 arc second pixel size (favored by many amateurs, whose CCD budgets don't run to as many pixels as the professionals'), J_{min} is 2.2×10^9, which corresponds to a surface brightness of -28 magnitudes per square arc second, which is how this is usually quoted in astronomical circles.[53]

If we were merely users of this instrument, that would be that. However, we're designing it from scratch, so we have complete flexibility to trade off the operating parameters. What should Ω be? For detection of faint extended objects such as nebulae, we set Ω as large as is consistent with adequate image resolution so that the photoelectrons dominate the dark current. For stellar objects, such as star clusters, we want Ω small, since if all the light from each star is going on one pixel anyway, shrinking Ω reduces the sky light while keeping the dark current and signal current the same. (There's no point in going too far – research-class telescopes keep Ω a few times smaller than the Airy disc.)

For a pixel size of 0.22 arc seconds, J_{min} is 1.1×10^{11}, which means that with a one hour exposure on a really great seeing night, we could detect a star of magnitude 25.6 with 3σ confidence. (We've left out the noise contributed by the calibration process, which may be significant if there aren't enough calibration frames – see Section 3.8.19.)

---

51 Optec Corp, *Model PFE–1 Technical Manual*.

52 Newberry, M.V. (1994). The signal to noise connection. *CCD Astronomy*, Summer (copyright ©1994 by Sky Publishing). Reprinted with permission.

53 This weird unit is a bit like quoting power spectral densities in dBm/Hz – convert back to watts before you multiply by the object size, or you'll get some very strange numbers (see Section 13.3.1).

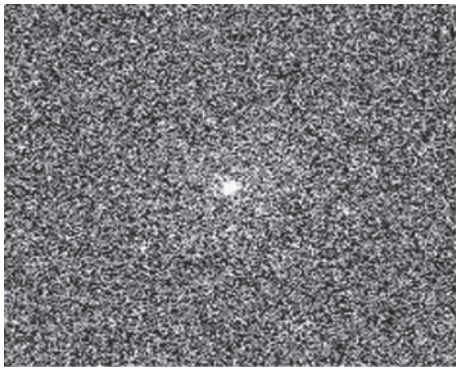


Figure 1.12 CCD image of the large spiral galaxy M100: 1 s integration.

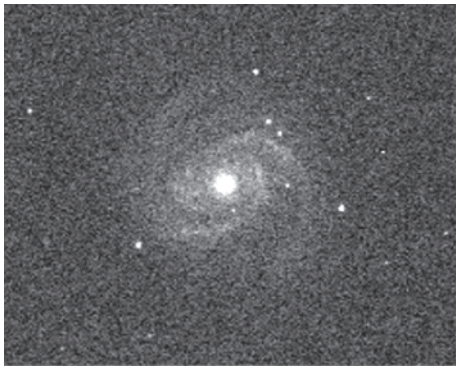


Figure 1.13 10 s integration.

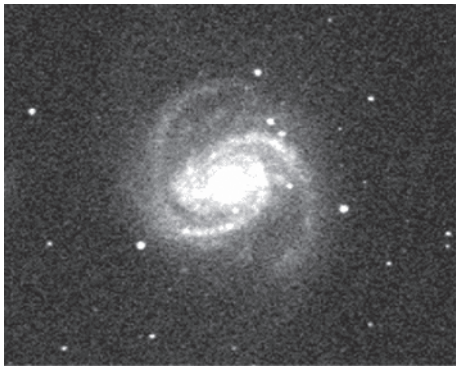


Figure 1.14 100 s integration.

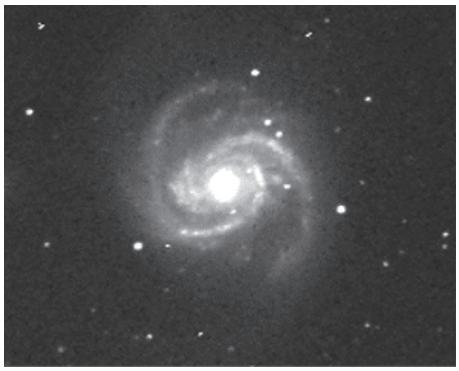


Figure 1.15 1000 s integration.

1.8 Signal Processing Strategy

Once the photocurrent leaves the detector, the work of the optical system is done and the maximum SNR has been fixed by the photon budget. The signal- processing job is to avoid screwing this up by needlessly rejecting or distorting the signal, on the one hand, or by admitting noise unnecessarily, on the other hand. Signal-processing operates under physical constraints having to do with causality and maximum allowable time delays and practical ones such as complexity and cost. Most of the time, the inherently highly parallel data coming in as light is converted to a serial electrical channel, so electrical wave filtering is appropriate.

1.8.1 Analog Signal Processing

The first bit of signal processing is the detector front-end amplifier, in which should be included any summing or subtracting of photocurrents, e.g. in a differential position measurement using a split detector. As we saw in Example 1.10, a bad front-end can hobble the system by adding noise and having too narrow a bandwidth with a given detector. Most designers are uncomfortable designing front-ends, and many wind up crippling themselves by buying a packaged solution that isn't appropriate to the problem; there's detailed help available in Chapters 3 and 18.

In a baseband (near DC) measurement, the next part is usually filtering and then digitizing. AC signals will usually need frequency conversion and detection too. The filter should roll off by at least 6N dB (where N is the number of bits) at the Nyquist limit of the digitizer, (see Section 17.4.4) and should reject strong spurious signals (e.g. 120 Hz from room lights) that would otherwise overwhelm subsequent stages or require more bits in the digitizer. In low-SNR situations, the bandwidth of the filter should be the same as that of the signal. Although some of the signal is lost, maximum SNR occurs when the filter's frequency response is identical with the signal's power spectrum. Filters wider than that pass more noise than necessary, and narrower ones clip off too much signal. Some filters have much better time responses than others. Have a look at the Bessel and equiripple group delay filters of Section 15.9.4. To get the best SNR with pulses, use a matched filter (Section 13.8.10).

We need to be able to convert time domain to frequency domain specifications. Remember that a 1-s averaging window corresponds to a bandwidth of 0.5 Hz at DC (Section 13.2.5). The reason is that in the analytic signal picture negative frequencies get folded over into positive ones. However, that same 1-s window and the same folding, applied to an AC measurement, gives 0.5 Hz both above and below the (positive frequency) carrier, a total of 1 Hz. The result is that a baseband (near-DC) measurement has half the noise bandwidth[54] of an AC measurement with the same temporal response. The resulting factor of 3 dB may be confusing.

The digitizing step requires care in making sure that the dynamic range is adequate; an attractively priced 8-bit digitizer may dominate the noise budget of the whole instrument, and the 12-bit one integrated in your microcontroller won't be a great deal better. (Almost all MCU ADCs are the pits – good enough for monitoring temperature sensors and power supplies, but not for your actual data.) For signals at least a few ADUs[55] in size, an ideal digitizer contributes additive *quantization noise* of $1/\sqrt{12}$ ADU to the signal, but real A/Ds may have as much as several ADUs of technical noise and artifacts (up to ~100 ADUs for $\Delta - \Sigma$ ADCs, see Sections 14.11.5 and 13.11.1), so check the data sheet. Converter performance generally degrades by 1–3 bits' worth between DC and the Nyquist frequency. Bits lost here cannot be recovered by postprocessing, so be careful to include a realistic model of digitizer performance in your conceptual design.[56]

1.8.2 Back-End Processing Strategy

With the current abundance of computing power, most measurements will include a fair bit of DSP. This may take the form of simple digital filtering to optimize the SNR and impulse response of the system, or may be much more complicated, such as digital phase-locked loops or maximum likelihood estimators of signal properties in the presence of statistically nonstationary noise and interference. In general, the difference between a simplistic back-end strategy and an optimal one is several decibels; this may seem trivial, but remember that a 3 dB improvement in your detection strategy can sometimes save you half your laser power or 2/3 of the weight of your collection optics.

---

54 Noise bandwidth is the equivalent width of the power spectrum of the filter (See Section 13.2.5.) If we put white noise into our filter, the noise bandwidth is the output power divided by the input noise power spectral density (in W/Hz), corrected for the passband insertion loss of the filter.

55 An ADU (analog-to-digital converter unit) is the amount of signal required to cause a change of 1 in the least significant bit of the converter. Often, confusingly called an LSB.

56 High-end oscilloscopes are a partial exception – they overcome timing skew and slew-dependent nonlinearity by calibrating the daylights out of themselves. They cost $100k and are only good to 6–7 bits.

1.8.3 Putting It All Together

Let's catch our breath for a moment. We began this chapter with the aim of learning how to do a complete feasibility calculation for an optical instrument, which we should now be fully equipped to do. We have covered a lot of ground in a short time, so don't be discouraged if it takes a while for it all to slot together. It becomes easier with practice, and standing around with one or two smart colleagues doing this stuff on a white board is about the most fun of any technical activity. To sum up, a conceptual design goes like this:

1. *Write down what you know.* Get a handle on the basic physics, figure out which is the largest effect operating, and estimate how much smaller the next biggest one is.
2. *Think up a measurement principle.* This may be routine, or may take a lot of imagination. Use it to take a rough cut at your photon budget. From that, estimate, e.g. the laser power required, the size of the detection area needed, and so on.
3. *Simplify the analysis.* Use limiting cases, but estimate where they will break down.
4. *Make a very rough optical design*. Choose the NAs, wavelength, working distances, and so on. Check that it makes physical and economic sense and won't be impossible to manufacture or (especially) to align.
5. *Guess a detection and signal-processing strategy.* One might choose a large germanium photodiode operating at 1 V reverse bias, followed by a somewhat noisy front end amplifier, a Butterworth lowpass filter, and a fixed threshold. Watch out for sources of systematic error and drift, e.g. etalon fringes or spectral changes in the source.
6. *Make a detailed photon budget.* See if your scheme will do the job, without unrealistically difficult steps.[57] If so, you have a measurement. If not, try step (5) again. If no amount of background reduction, low-noise circuitry or signal processing will help, go back to (2) and think harder. It's often very worthwhile going round that loop a few times before deciding.
7. *Try it the other way round.* Horace Darwin (grandson of the famous natural philosopher) was a fine instrument builder. One of his maxims was "Always try it the other way round." Perhaps it would be easier to scan by moving the sample instead of the beam, or to measure extinction instead of scattering, or to modulate the sample rather than the laser. The decisions you're making now will be ones you'll have to live with for awhile, so turn it over in your mind a few times before committing. We'll talk more about this in Section 11.7.
8. *Check it for blunders.* Do it all over again a different way, and use scaling arguments to make sure the power laws are all correct. If you go ahead with this instrument, a lot will be riding on your choice of approach and the correctness of your calculation – don't scrimp on time and effort here. Try checking the scaling with parameter values and see if you get the right behavior.[58] If you have one or two colleagues who are friendly but difficult to convince, try it out on them, to see if they can poke any holes in your logic.

Remember that what will get you is not misplacing a factor of Boltzmann's constant – that's large and obvious – but rather the factors of order 1 that you haven't thought out carefully. These are things like using the peak value when the RMS is what's relevant, or forgetting that when you multiply two similar peaks together the result is $\sqrt{2}$ times narrower, or assuming that the bandwidth of a 1-s averaging window is 1 Hz (see Section 13.2.5). A few of these taken together can put you off by factors of 10 or more – very embarrassing.

Estimating and controlling systematic errors is one of the high arts of the designer, since they don't obey any nice theorems as random errors sometimes do. For now we'll just try to keep the optical system simple and the processing modest and linear. Getting too far away from the measurement data is a ticket to perdition.

Let's do an extended example that illustrates many of the concepts of this chapter.

Example 1.12 ***Quantum-limited Scanning Interferometer Design Concept***

The ISICL (*In Situ* Coherent Lidar) system[59] detects submicron (> 0.25 μm) contaminant particles in plasma chambers and other semiconductor processing equipment. As shown in Figure 1.16, it works by scanning a weakly focused laser beam around inside the chamber through a single window, and detecting the light backscattered from particles.

Backscatter operation is very difficult. Between the strong plasma glow and the stray laser light bouncing off the back wall of the chamber, it must deal with very bright stray light – thousands of times worse than that seen by an ordinary dark-field detector, and around a million times brighter than a nominally detectable particle. Coherent detection is probably the

57 As in the light bulb spectrometer of Example 17.10.

58 This would have saved the author some heartburn on one occasion – in a long computation in Mathcad for a fiber-coupled spectrometer, he got two function parameters interchanged. Unfortunately, they were the fiber NA and the number of fibers; 0.27 fibers of NA = 15 would give a lot more light than 15 fibers of NA 0.27 (In fact 55 times as much because the étendue goes as NA^2). The instrument worked anyway, but it was a harder job than anticipated.

59 Hobbs, P.C.D. (1995). ISICL: *In situ* coherent lidar for particle detection in semiconductor processing equipment. *Appl. Opt.* 34 (9): 1579–1590. https://electrooptical.net/www/isicl/isiclAO.pdf.

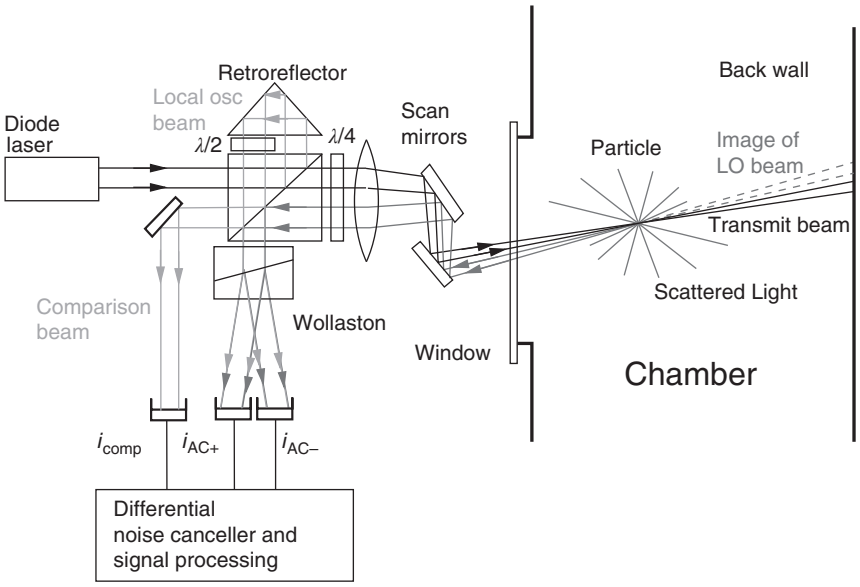


Figure 1.16 The ISICL sensor is an alignment-insensitive scanning interferometer for finding submicron particles in hostile places such as plasma chambers. The corner cube has the same optical properties as a lens with a scatterer near focus, so the alignment is trivial. The $\lambda/4$ plate makes the RX beam *s*-polarized so that it gets reflected by the beam splitter, and the $\lambda/2$ plate allows the polarization of the LO beam to be adjusted to compensate for the polarization rotation in the corner cube and to adjust the relative strengths of the LO and compensation beams.

only feasible measurement strategy. With a coherent detector, a particle crossing near the focus of the beam gives rise to a short ($\gtrsim$3 μs) tone burst, whose duration is the transit time of the beam across the particle and whose carrier frequency is the Doppler shift from its radial motion. These tone bursts are filtered and amplified, then compared with a threshold to determine whether a particle event has occurred. Since the phase of the tone burst is random, the highest peak may be positive or negative, so bipolar thresholds are used.

The sensitivity of an instrument is expressed by the minimum number of photons it requires in order to detect a particle, which is a function of the confidence level required. A false alarm rate (FAR) of 10^{-11} in the measurement time corresponds to about 1 false count per day with bipolar thresholds in a 1 MHz bandwidth. From Section 13.6.19, the false count rate depends only on the measurement bandwidth and the ratio α of the threshold to the RMS noise voltage. A false count rate of 10^{-11} per inverse bandwidth requires $\alpha = 7.1$.

In an AC measurement, the shot noise level is equivalent to a single coherently detected noise photon in the measurement time, so it might seem that seven scattered photons would be enough. Coherent detectors detect the amplitude rather than the intensity of the received light, though, so to get a peak signal current 7.1 times the RMS noise current, we actually need $7.1^2 \approx 50$ photons per burst.

We have a first estimate of how many photons we need, so let's look at how many we expect to get. For a particle in the center of the sensitive region, the received power is the product of the transmit beam flux times the differential scattering cross section $\partial\sigma/\partial\Omega$ of the particle, times the detector projected solid angle Ω_d. Working in units of photons is convenient because the relationship between photon counts and shot noise means that electrical SNR is numerically equal to the received photon count per measurement time. This works independently of the signal and filter bandwidth. Initially, we ignore the losses imposed by the matched filter.

For a Gaussian transmit beam of power P at wavelength λ, focused at a numerical aperture NA, Table 1.1 gives the photon flux at the beam waist as

$$F(P, \lambda, \mathrm{NA}) = \frac{2\pi(\mathrm{NA})^2 P\lambda}{hc}. \tag{1.94}$$

Assuming the scattered field is constant over the detector aperture, Table 1.1 predicts that the effective detector solid angle is

$$\Omega_d = \pi(\mathrm{NA})^2 \tag{1.95}$$

and so the expected number of photons available per second is

$$\langle n_0 \rangle = \frac{2\pi^2(\mathrm{NA})^4 P\lambda}{hc}\frac{\partial\sigma}{\partial\Omega}. \tag{1.96}$$

Not all of these will be detected owing to imperfect efficiency of the optics and the detector. The quantum efficiency η of the detector is the average number of detection events per incident photon on the detector; it's always between 0 and 1. (Photodetectors generally give one photoelectron per detection event, so $\eta \langle n_0 \rangle$ is the number of photoelectrons before any amplification.) A good quality, antireflection coated silicon photodiode can have quantum efficiencies $\eta \approx 0.95$ in the red and near IR, but there are in addition front surface reflections from windows, and other optical losses. A receiver whose end-to-end efficiency is over 0.5 is respectable, and anything over 0.8 is very good. A value of 0.9 can be achieved in systems without too many elements, but not cheaply. (We should also multiply by the square of the Strehl ratio to account for aberrations; see Example 9.6 and Section 9.5.4.)

The SNR can be traded off against measurement speed by the choice of scanning parameters and filter bandwidths. Narrowing the bandwidth improves the time-averaged SNR but smears pulses out. In a pulsed measurement, the optimal filter is the one that maximizes the received SNR at the peak of the pulse. For pulses in additive white Gaussian noise, the optimum filter transfer function is the complex conjugate of the received pulse spectrum (this filter is not necessarily the best in real life, but it's an excellent place to start – see Section 13.8.10). Such a matched filter imposes a 3 dB signal loss on a Gaussian pulse.[60] However, the measurement detects threshold crossings, and for the weakest detectable signals, the peaks just cross the threshold. Thus, this 3 dB is made up by the factor of $\sqrt{2}$ voltage gain from the peak-to-RMS ratio so that the minimum detectable number of photons (in the deterministic approximation) is still α^2.

We can estimate the stray light in the following way. Assume a detector requiring 50 photons for reliable particle identification, operating in backscatter with a 100-mW laser at 820 nm and NA = 0.008. Assume further (optimistically) that the back wall of the chamber is a perfectly diffuse (Lambertian) reflector so that the light is scattered into π sr. The incident beam has 4×10^{17} photons/s so that the backscattered stray light has average brightness 1.3×10^{17} photons/s/sr; the detector solid angle is $\pi(\text{NA})^2 \approx 0.0002$ sr, so the total detected signal due to stray light is about 2.6×10^{13} photons/s. Assuming that a particle takes 3 μs to yield 50 photons (1.7×10^7 photons/s), the stray light is 10^6 times more intense than the signal from a nominally detectable particle. What is worse, the signal from the back wall exhibits speckles, which move rapidly as the beam is scanned, giving rise to large (order-unity) fluctuations about the average stray light intensity. The angular size of the speckles (and hence the bandwidth of the speckle noise) depends on the distance from the focus to the chamber wall, and on the surface finish of the wall. Peak background signals are generally much larger than this average.

The Doppler frequency shift in the detected signal due to a particle traveling with velocity $\mathbf{v}$, encountering incident light with wave vector $\mathbf{k}_i$ and scattering it into a wave with wave vector $\mathbf{k}_s$ is

$$f_d = \mathbf{v} \cdot (\mathbf{k}_s - \mathbf{k}_i)/(2\pi). \tag{1.97}$$

For a system operating in backscatter, $\mathbf{k}_s - \mathbf{k}_i \approx -2\mathbf{k}_i$. At 830 nm, a particle moving axially at 50 cm/s will give rise to a tone burst whose carrier frequency is about 1.22 MHz. This is the nominal maximum particle axial velocity for the original ISICL. (A subsequent version was designed to detect 0.15 μm metal particles at velocities up to 3 km/s (Mach 9) – see Example 11.3.)

The remaining engineering problems center on the choice of detection bandwidths and the accurate estimation of the shot noise level in the presence of large signals due to particles. In the present case, where the Doppler shift may be large compared to the transit time bandwidth, the peak frequency of the received spectrum of a given pulse is not known in advance. The optimal signal processing system depends on the range of particle velocities expected. In quiet plasma chambers, where most particles orbit slowly within well-defined traps, the maximum expected velocity may be as low as 5 cm/s, whereas in an environment such as a rapidly stirred fluid tank or the roughing line of a vacuum pump, the velocity range may be much greater. The scan speed of the beam focus through the inspected volume is much higher (about 20 m/s).

With carrier frequencies ranging from 0 to 1.22 MHz, the Doppler bandwidth is much larger than the transit time bandwidth (150 kHz for a 3 μs burst FWHM), so that it is inefficient to perform the thresholding operation in a single band. In the present system, four bands are used to cover the Doppler bandwidth. This is frequently best in low-NA systems like ISICL, where the focus is many wavelengths across.

In a thresholding operation, it is essential to set a high enough threshold that the sensor does not report erroneously high particle counts, possibly resulting in needless down time for the processing tool being monitored. At the same time, it is economically important to use the available laser power as efficiently as possible; at the time, the laser used in this sensor cost over $1200, so that (loosely speaking) too-high thresholds cost $300 per decibel. The signal-processing strategy is to set separate bipolar thresholds for each band, using an automatic thresholding circuit. This circuit exploits the noise

60 Skolnik, M. (ed.) (1990). *Radar Handbook*, 2e, 3.21–3.23. New York: McGraw-Hill.

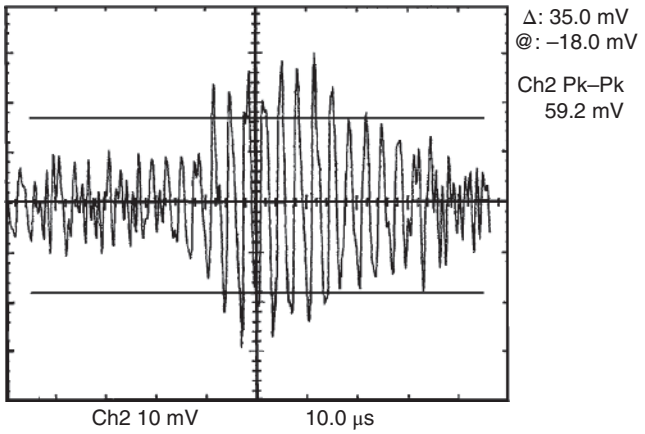


Figure 1.17 Measured tone burst caused by a 0.8 μm polystyrene latex (PSL) sphere crossing the beam near focus. The horizontal cursors are the predicted peak–peak value of the output. The bandwidth of the signal shown is several times wider than the actual measurement bandwidth, which prevents distortion of the tone burst but increases the noise background.

canceller's accurately known noise amplitude statistics to servo on the false counts themselves and ignore signal pulses, however large they may be.[61] This is superficially similar to a CFAR (constant false alarm rate) servo as used in radar; the technologies employed are quite different, however, since a radar system can look at a given target many times, and its noise is very non-Gaussian. The ISICL FAR tracker can accurately servo the FAR at a level much below the true count rate in most applications.

Figure 1.17 shows a typical tone burst from a 0.8 μmø PSL sphere, together with cursors showing the predicted peak-to-peak voltage for a particle passing through the center of the sensing volume. For a particle exactly in focus, the photon flux predicted by (1.96) is converted to a signal current i_{AC} by (1.73), and to a voltage by multiplying by the known current to voltage gain (*transimpedance*) of the front-end and any subsequent amplifiers. Because of the known relationship between the signal size and the shot noise in a coherent detector, we can check the ratio of the RMS noise to the signal easily as well (the aberration contribution is calculated in Example 9.6). The measured system parameters are NA = 0.0045, $P = 90$ mW, $\lambda = 820$ nm, $\eta = 0.64$. Taking into account the addition of the noise and signal, the error is less than 20% (1.6 dB), indicating that the theory correctly predicts the signal size and SNR.

61 See U. S. Patent 5,204,631.