FABIEN BARON & STUART JEFFERIES

HIGH ANGULAR RESOLUTION IMAGING FOR ASTRONOMY

ADAPTIVE OPTICS, IMAGE RESTORATION & INTERFEROMETRY

OPEN ASTROPHYSICS BOOKSHELF

About the cover: In mid-August 2010 European Southern Observatory's Photo Ambassador Yuri Beletsky snapped this photo at ESO's Paranal Observatory, Chile. A group of astronomers were observing the center of the Milky Way using the laser guide star facility at Yepun, one of the four Unit Telescopes of the Very Large Telescope (VLT). Yepun's laser beam crosses the southern sky and creates an artificial star at an altitude of 90 km high in the Earth's mesosphere. The Laser Guide Star (LGS) is part of the VLT's adaptive optics system and is used as a reference to correct the blurring effect of the atmosphere on images. The color of the laser is precisely tuned to energize a layer of sodium atoms found in one of the upper layers of the atmosphere – one can recognize the familiar color of sodium street lamps in the color of the laser. This layer of sodium atoms is thought to be a leftover from meteorites entering the Earth's atmosphere. When excited by the light from the laser, the atoms start glowing, forming a small bright spot that can be used as an artificial reference star for the adaptive optics. Using this technique, astronomers can obtain sharper observations. For example, when looking towards the center of our Milky Way, researchers can better monitor the galactic core, where a central supermassive black hole, surrounded by closely orbiting stars, is swallowing gas and dust. Taken with a wide angle lens, this photo covers about 180° of the sky.
Credit & Copyright: ESO/Yuri Beletsky, CC BY 4.0

© 2017 Fabien Baron & Stuart Jefferies
git version c2f32663 . . .

🅭🅯🄎
Except where explicitly noted, this work is licensed under the Creative Commons Attribution-ShareAlike 4.0 International (CC BY-SA 4.0) license.

Preface

These notes are currently being written while teaching a graduate-level astronomy course on adaptive optics, image restoration and optical interferometry at Georgia State University. The text layout uses the `tufte-book` (`https://tufte-latex.github.io/tufte-latex/`) LaTeX class: the main feature is a large right margin in which the students can take notes; this margin also holds small figures and sidenotes. Exercises are embedded throughout the text. These range from "reading exercises" to longer, more challenging problems. In addition, there are several numerical exercises in Python and Julia (available from `https://julialang.org`). Because the exercises are spread throughout the text, there is a "List of Exercises" in the front matter to help with looking for specific problems.

Please be advised that these notes are under active development; to refer to a specific version, use the eight-character stamp labeled "git version" on the copyright page.

Contents

1 Fourier Optics 1
- 1.1 Characterizing instrument aberrations 1
  - Point source 1
  - The pupil field 1
  - Optical path length 2
  - Zernike polynomials 3
  - Point spread function 5
  - Airy pattern 6
  - Polychromatic PSF 7
- 1.2 Image formation 8
  - Imaging is convolving 8
  - Fast convolution 8
  - Optical transfer function 8
- 1.3 Sparse aperture characteristics and OTF 9
- 1.4 Some practical matters 11
  - Convolution and correlation 11
  - Autocorrelation 11
  - Computing convolution and correlation using the discrete Fourier Transform 12
  - Image sampling 12
- 1.5 Atmosphere and turbulence 13
  - The Kolmogorov model of turbulence 14
  - C_n^2 profile 16
  - Fried parameter: r_0 16

Atmospheric time constant: τ_0 17
Isoplanatic angle: θ_0 18
Layer model for atmosphere 18
Scintillation 19
1.6 Imaging through atmospheric turbulence 21
Long-exposure images 22
Atmospheric transfer function 23
Speckle imaging 23
Shift-and-add 23
Speckle Interferometry 24

2 Image Restoration 27
Image preprocessing 27
Multiplicative deconvolution 28
Blind deconvolution 29
Diversity imaging methods 31

3 Adaptive Optics 33
3.1 Wavefront references, wavefront sensing 33
3.2 AO & MCAO 33
3.3 Image restoration with AO compensated data 33

4 Optical interferometry 35
4.1 Fringes 35
4.2 Visibilities 35
Radio versus optical 35
4.3 Fringe tracking 35
4.4 Aperture synthesis 35
4.5 Interferometric imaging 35

1

Fourier Optics

Fourier optics provides a description of the propagation of light based on harmonic analysis (the Fourier transform) and linear systems.

1.1 Characterizing instrument aberrations

Point source

A point source is an object whose angular size is below the angular resolution of the instrument. It is said to be *unresolved*. Mathematically, a point source is a 2D delta function.

$$\delta(x,y) = \begin{cases} 1 & \text{if } (x,y) = (x_0, y_0) \\ 0 & \text{elsewhere} \end{cases} \tag{1.1}$$

The pupil field

The pupil function describes how a light wave is affected upon transmission through an optical imaging system. More specifically, it is a two-dimensional complex function of the position in the pupil that indicates the relative change in amplitude and phase of the light wave. Imperfections in the optics typically have a direct effect on the pupil function, it is therefore an important tool to study optical imaging systems and their performance. The pupil function $P(u,v)$ can be split into its modulus and phase:

$$P(u,v) = A(u,v) \exp i\Phi(u,v), \tag{1.2}$$

where the modulus $A(u,v)$ includes the aperture function (percentage of light blocked) and throughput losses, and the phase $\Phi(u,v)$ describes the aberrations of the system. The pupil function of an ideal system (no aberrations, 100% transmission) is equal to one at every point within the pupil, and zero outside of it. In the case of a circular pupil of unit radius, the pupil function is $P(u,v) = \Pi(u,v)$, where:

$$\Pi(u,v) = \begin{cases} 1 & \forall\,(u,v) | \sqrt{u^2+v^2} \leq 1 \\ 0 & \text{elsewhere} \end{cases} . \tag{1.3}$$

The circular pupil is the most common type of aperture in astronomy. However, the continuing quest for finer angular resolution demands larger and larger apertures. Unfortunately, the manufacturing cost for monolithic optics increases faster than the diameter of the optic squared, and we are already close to the limit of what is financially feasible. Consequently, for applications that do not require extremely high sensitivity, the notion of a sparse aperture offers a novel solution for decreasing the volume and weight of a telescope while allowing for large pupil sizes. Golay[1] proposed that there is an optimal imaging configuration for sparse arrays that depends on the number of sub-apertures.

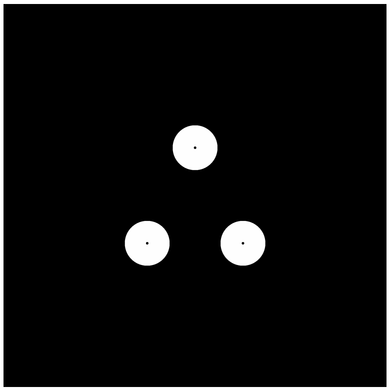

Figure 1.1: A Golay array with three sub-apertures.

Optical path length

The optical path length (OPL) or optical distance is the product of the geometric length of the path light through the system, by the index of refraction of the medium through which it propagates. A difference in optical path length between two paths is often called the optical path difference (OPD). Optical path length is important because it determines the phase of the light and governs interference and diffraction of light as it propagates.

The OPD corresponds to the phase shift undergone by the light emitted from two previously coherent sources when passed through mediums of different refractive indices. For example, a wave passed through glass will appear to travel a greater distance than an identical wave in air. This is because the source in the glass will have experienced a greater number of wavelengths due to the higher refractive index of the glass. The OPD can be calculated from the following equation:

$$\mathrm{OPD} = d_1 n_1 - d_2 n_2 \tag{1.4}$$

where d_1 and d_2 are the distances of the ray passing through medium 1 or 2, n_1 and n_2 the refractive indices. In a medium of constant refractive index n the OPL for a path of physical length d is just $\mathrm{OPL} = nd$. If the refractive index varies along the path, the OPL is given by

$$\mathrm{OPL} = \int_C n(s)\mathrm{d}s, \tag{1.5}$$

where $n(s)$ is the local refractive index as a function of distance, s, along the path C. Finally, the phase shift introduced by an OPD is

[1] M. Golay, "Point Arrays Having Compact Non-redundant Autocorrelations", J.Opt Soc. Am., vol. 61, pp 272, 1971.

given by:

$$\Phi = \frac{2\pi}{\lambda}\text{OPD} \quad \text{(rad)} \tag{1.6}$$

provided the OPD and λ are expressed in the same units. Consequently, when considering e.g. $1/2$ wavelength of piston aberration, we are talking about a phase shift of $\Phi = \frac{2\pi}{\lambda} \times \frac{1}{2}\lambda = \pi$ radians.

Zernike polynomials

The Zernike polynomials are orthogonal two-dimensional functions, which form a basis defined over the circular support area; typically the pupil planes in classical optical imaging at visible and infrared wavelengths through systems of lenses and mirrors of finite diameter. They were introduced by Frits Zernike in the 1930s to model the pupil wavefront. [2]

A circular wavefront profile Φ associated with aberrations may be mathematically modeled using a linear sum of Zernike polynomials.

$$\Phi = \sum_{j=1}^{N} a_j Z_j \tag{1.7}$$

where j is the main index (Noll index) of Zernike polynomials Z_j, and a_j its scalar strength in the wavefront.

These Zernike coefficients are linearly independent, thus individual aberration contributions to an overall wavefront may be isolated and quantified separately. Hence an aberrated wavefront profile may be decomposed into Zernike polynomials since:

$$\forall i, \quad Z_i \cdot \Phi = \sum_{j=1}^{N} a_j Z_i \cdot Z_j = a_i Z_i \cdot Z_i = \pi a_i \tag{1.8}$$

where it is assumed $Z_i \cdot Z_j = \pi\delta_{ij}$ (π is often chosen for normalization).

Going beyond simple indexing, there are even and odd Zernike polynomials. The even Zernike polynomials are defined as:

$$Z_n^m(\rho,\theta) = R_n^m(\rho)\,\cos(m\,\theta) \tag{1.9}$$

and the odd Zernike polynomials as

$$Z_n^{-m}(\rho,\theta) = R_n^m(\rho)\,\sin(m\,\theta) \tag{1.10}$$

where m and n are nonnegative integers with $n \geq m$, θ is the azimuthal angle, ρ is the radial distance $0 \leq \rho \leq 1$, and R_n^m are the radial polynomials defined below. Zernike polynomials have the property of being limited to a range of -1 to $+1$, i.e. $|Z_n^m(\rho,\theta)| \leq 1$. The radial polynomials $R_n^m$are defined as

$$R_n^m(\rho) = \sum_{k=0}^{\frac{n-m}{2}} \frac{(-1)^k\,(n-k)!}{k!\,\left(\frac{n+m}{2}-k\right)!\,\left(\frac{n-m}{2}-k\right)!}\,\rho^{n-2\,k} \tag{1.11}$$

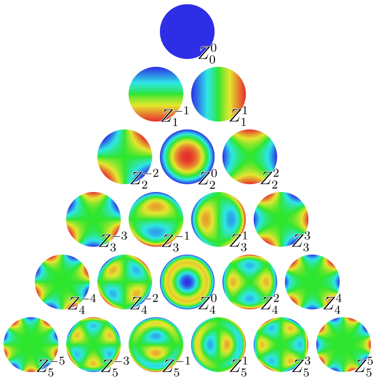

Figure 1.2: The first 21 Zernike polynomials, ordered vertically by radial degree and horizontally by azimuthal degree. Each represent a possible mode of the pupil phase.

---

[2] At this time, the perfect point spread function in the presence of diffraction had already been described by George Biddell Airy, as early as 1835. It took almost hundred years to arrive at a comprehensive theory and modeling of the point image of aberrated systems (Zernike and Nijboer). The analysis by Nijboer and Zernike describes the intensity distribution close to the optimum focal plane. Recent analytic results have made it possible to extend Nijboer and Zernike's approach for point-spread function evaluation to a large volume around the optimum focal point. This extended Nijboer-Zernike (ENZ) theory is instrumental in studying the imperfect imaging of three-dimensional objects in confocal microscopy or astronomy under non-ideal imaging conditions. The ENZ-theory has also been applied to the characterization of optical instruments with respect to their aberration by measuring the through-focus intensity distribution and solving an appropriate inverse problem.

Noll index j	Radial degree n	Azimuthal degree m	Z_j	Classical name
1	0	0	1	Piston
2	1	1	$2\rho\cos\theta$	Tip (X-Tilt)
3	1	-1	$2\rho\sin\theta$	Tilt (Y-Tilt)
4	2	0	$\sqrt{3}(2\rho^2-1)$	Defocus (longitudinal position)
5	2	-2	$\sqrt{6}\rho^2\sin 2\theta$	Oblique astigmatism
6	2	2	$\sqrt{6}\rho^2\cos 2\theta$	Vertical astigmatism
7	3	-1	$\sqrt{8}(3\rho^3-2\rho)\sin\theta$	Vertical coma
8	3	1	$\sqrt{8}(3\rho^3-2\rho)\cos\theta$	Horizontal coma
9	3	-3	$\sqrt{8}\rho^3\sin 3\theta$	Vertical trefoil
10	3	3	$\sqrt{8}\rho^3\cos 3\theta$	Oblique trefoil
11	4	0	$\sqrt{5}(6\rho^4-6\rho^2+1)$	Primary spherical
12	4	2	$\sqrt{10}(4\rho^4-3\rho^2)\cos 2\theta$	Vertical secondary astigmatism
13	4	-2	$\sqrt{10}(4\rho^4-3\rho^2)\sin 2\theta$	Oblique secondary astigmatism
14	4	4	$\sqrt{10}\rho^4\cos 4\theta$	Vertical quadrafoil
15	4	-4	$\sqrt{10}\rho^4\sin 4\theta$	Oblique quadrafoil

Table 1.1: The first few Zernike modes, ordered by Noll index j, normalized so that $\int_0^{2\pi}\int_0^1 Z_j^2\,\rho\,d\rho\,d\theta = \pi$.

for $n-m$ even, and are identically 0 for $n-m$ odd.

The first few Zernike polynomials are show in Table 1.1; ρ is the normalized pupil radius with $0 \leq \rho \leq 1$, θ is the azimuthal angle around the pupil with $0 \leq \theta \leq 2\pi$, and the fitting coefficients $a_0, \ldots, a_8$ are the wavefront errors in wavelengths.

Their advantages are the simple analytical properties inherited from the simplicity of the radial functions and the factorization in radial and azimuthal functions; this leads, for example, to closed-form expressions of the two-dimensional Fourier transform in terms of Bessel functions. They are commonly used in adaptive optics, where they can be used to effectively cancel out atmospheric turbulence. For example, one of the Zernike terms (for m = 0, n = 2) is called defocus. By coupling the output from this term to a control system, an automatic focus can be implemented. Another advantage is that it is often useful to describe the wave front aberrations by the rms value σ averaged over the pupil. If the wave front is modeled as a Zernike series then

$$\sigma = \sqrt{\sum_{j=2}^{\infty} a_j^2} \tag{1.12}$$

Their disadvantage, in particular if high n are involved, is the unequal distribution of nodal lines over the unit disk, which introduces ringing effects near the perimeter ($\rho \approx 1$), which often leads attempts to define other orthogonal functions over the circular disk (e.g., disk harmonics). Wavefronts with very steep gradients or very high spatial frequency structure, such as produced by wave propagation through atmospheric turbulence, are not well modeled by Zernike polynomials, which tend to low-pass filter the fine three-dimensional spatial definition in the wavefront. In this case, other fitting methods such as fractals or singular value decomposition may yield improved fitting results.

EXERCISE 1.1—

1. Compute the Hadamard product of the first 20 Zernikes using the library provided with this book, and display the resulting 20×20 triangular array. How is the orthogonality affected on a non-circular pupil (e.g. annular) ?
2. Perform the Zernike decomposition of the provided wavefront. Given the results of the previous exercise, what would happen if attempting the same for an annular pupil ?

Point spread function

A point of light in the object plane is spread out in the image plane according to the imaging system's impulse response function. If the illumination is spatially incoherent then the impulse response is called the point spread function (PSF), i.e. the PSF is the image of an unresolved object through the system. The PSF can also be understood as the power-spectrum of the normalized pupil field function, i.e.,

$$PSF = \left| \mathscr{F}^{-1}(P_n) \right|^2 \tag{1.13}$$

where

$$P_n(u,v) = A_n(u,v) \exp i\Phi(u,v), \tag{1.14}$$

and

$$A_n(u,v) = \frac{A(u,v)}{\sqrt{\sum\limits_{u,v} A^2(u,v)}} = \frac{|P(u,v)|}{\sqrt{\sum\limits_{u,v} |P(u,v)|^2}}. \tag{1.15}$$

This normalization ensures that the PSF has unit volume.
A PSF has a number of other noteworthy characteristics:

1. For a circular aperture of diameter D the first zero in the PSF's diffraction intensity pattern occurs at $1.22\lambda/D$. This corresponds to the angular separation at which two point sources become resolved and is known as the Rayleigh resolution limit.[3]

2. The width of the PSF at half its maximum intensity is called the full width at half maximum (FWHM). For a circular aperture the FWHM is given by $1.03\lambda/D$. Note, the peak of a Gaussian profile has a similar shape to that of the Airy function and for a standard deviation $\sigma_G = 0.44\lambda/D$, has the same FWHM.

3. The ratio of the peak intensity of the PSF to the peak intensity of the unaberrated PSF (i.e. the PSF you would get if the aberration phase $\Phi = 0$) is called the Strehl ratio S:

$$S = \frac{\max\{\mathrm{PSF}(\Phi)\}}{\max\{\mathrm{PSF}(\Phi = 0)\}} \leq 1 \tag{1.16}$$

---

[3] The angular resolution in arcseconds is given by $\theta = 0.25\lambda/D$, where λ is in microns and D is in meters.

This ratio provides a simple meaningful way of expressing the effect of wave front aberrations on image quality: the higher the Strehl ratio, the better the image resolution. $S = 1$ corresponds to a diffraction-limited image (the best). If σ – the *rms* value averaged over the wavefront Φ – is known, then we can use the Maréchal formula to approximate the Strehl: $S \simeq \exp(-\sigma^2)$.

4. The PSF integral over the circle of radius β is called the encircled energy. This characteristic is important for observations of faint objects, where one wants to concentrate the photons in as small an area as possible.

Airy pattern

The Airy disc (or Airy pattern) is the best focused spot of light that a perfect lens with a circular aperture can make, limited only by the diffraction of light. Given by the squared modulus of the Fourier transform of the circular aperture Π:

$$I(\theta) = I_0 \left(\frac{2J_1(\frac{2\pi}{\lambda} R \sin\theta)}{\frac{2\pi}{\lambda} R \sin\theta} \right)^2 = I_0 \left(\frac{2J_1(x)}{x} \right)^2 \tag{1.17}$$

where I_0 is the maximum intensity of the pattern at the Airy disc center, J_1 is the Bessel function of the first kind of order one, R is the radius of the aperture, and θ is the angle of observation, i.e. the angle between the axis of the circular aperture and the line between aperture center and observation point. Expressed in linear units, $x = \frac{2\pi R}{\lambda} \frac{r}{d}$, where r is the radial distance from the optics axis in the observation plane and d the observation distance. Note that the limit for $\theta \to 0$ (or for $x \to 0$) is $I(0) = I_0$. We can now check the results from the previous section: the zeros of $J_1(x)$ are at $x \approx 3.8317, 7.0156, 10.1735, 13.3237, \ldots$. From this, it follows that the first dark ring in the diffraction pattern occurs where $\frac{2\pi R}{\lambda} \sin\theta = 3.8317$, or $\sin\theta \approx 1.22 \frac{\lambda}{D}$. The next dark rings are at $2.23 \frac{\lambda}{D}$ and $3.24 \frac{\lambda}{D}$. The half maximum of the central Airy disk (where $J_1(x) = x/2\sqrt{2}$) occurs at $x \approx 1.61633$, leading to the FWHM at $1.028 \frac{\lambda}{D}$. The maximum of the first ring occurs at $x \approx 5.13562$, or $1.63 \frac{\lambda}{D}$.

The intensity I_0 at the center of the diffraction pattern is related to the total power P_0 incident on the aperture by:

$$I_0 = \frac{E_A^2 R^2}{2d^2} = \frac{P_0 A}{\lambda^2 d^2} \tag{1.18}$$

where E is the source strength per unit area at the aperture, A is the area of the aperture ($A = \pi R^2$) and d is the distance from the aperture. The intensity at the maximum of the first ring is about 1.75% of the intensity at the center of the Airy disk. The expression for $I(\theta)$ above can be integrated to give the total power contained in the diffraction pattern within a circle of given size:

$$P(\theta) = P_0 \left[1 - J_0^2 \left(\frac{2\pi}{\lambda} R \sin\theta \right) - J_1^2 \left(\frac{2\pi}{\lambda} R \sin\theta \right) \right] \tag{1.19}$$

where J_0 and J_1 are Bessel functions. Hence the fractions of the total power contained within the first, second, and third dark rings (where $J_1(\frac{2\pi}{\lambda} R \sin\theta) = 0$) are 83.8%, 91.0%, and 93.8% respectively.

[4]

Polychromatic PSF

Equations 1.13 and 1.14 define the PSF for monochromatic light. However, many observations in astronomy are taken through filters with appreciable spectral bandwidth. In this case we need to use the polychromatic point spread function (PPSF), which is obtained by integrating the monochromatic intensities along the sampled spectrum taking into account both the spectral responses of the detector and any filters used, and (ideally) the spectral emission of the source. Practically, this is done by computing the PSF at a series of discrete wavelengths (color planes) and summing across all color planes weighted by the spectral response function, $S(\lambda)$:

$$PPSF = \frac{\sum_j S(\lambda_j) PSF(\lambda_j)}{\sum_j S(\lambda_j)}. \tag{1.20}$$

EXERCISE 1.2— Generate PSFs for:

1. Circular apertures of radii $R_1 < R_2 < R_3$.
2. Circular apertures with a secondary hole.
3. Circular apertures with a secondary hole and a support spider.
4. A sparse aperture called the Golay-3 pupil, with the following subaperture center (x, y) coordinates: $\left(-\frac{1}{2}, -\frac{\sqrt{3}}{6}\right)$, $\left(\frac{1}{2}, -\frac{\sqrt{3}}{6}\right)$ and $\left(0, \frac{\sqrt{3}}{3}\right)$.
5. A sparse aperture called the Golay-6 pupil, with the following subaperture center (x, y) coordinates: $\left(1, \frac{\sqrt{3}}{3}\right)$, $\left(\frac{3}{2}, -\frac{\sqrt{3}}{6}\right)$, $\left(0, -\frac{2\sqrt{3}}{3}\right)$, $\left(-1, -\frac{2\sqrt{3}}{3}\right)$, $\left(-1, \frac{\sqrt{3}}{3}\right)$, $\left(-\frac{1}{2}, \frac{5\sqrt{3}}{6}\right)$.
6. Look at different PSFs for the first 12 Zernikes.
7. Calculate Strehl for the previous PSFs. Be mindful of the normalization of the unaberrated pupil and of the normalization of the Fourier transform functional.

---

[4] IMAGE: 3 DIFFERENT APERTURES (CIRCULAR, SQUARE AND HEXAGONAL) AND THEIR CORRESPONDING PSFS.
IMAGE: IMAGE OF POINT SOURCE WITH: NO ABERRATIONS, 1 WAVE PTV SPHERICAL, 1 WAVE PTV ASTIGMATISM, 1 WAVE PTV SPHERICAL+ASTIGMATISM+COMA

1.2 Image formation

Imaging is convolving

By virtue of the linearity property of optical imaging systems, the image of the superposition of two objects O_1 and O_2 is the superposition of the two images:

$$I(O_1 + O_2) = I(O_1) + I(O_2). \tag{1.21}$$

Since the object-plane field is itself a weighted sum over 2D impulse response functions[5],

$$O(x_0, y_0) = \sum_{(u,v)} O(u,v)\ \delta(x_0 - u, y_0 - v) \tag{1.22}$$

the image plane field is the weighted sum over the images of these impulse response functions,

$$I(x_i, y_i) = \sum_{(u,v)} O(u,v)\ \mathrm{PSF}(x_i - u, y_i - v) \tag{1.23}$$

With the following definition for the discrete 2D convolution:

$$\begin{aligned}(f * g)[n,m] &= \sum_p \sum_q f[p,q]\, g[n-p, m-q] \\ &= \sum_p \sum_q f[n-p, m-q]\, g[p,q].\end{aligned} \tag{1.24}$$

we realize the image is given as the convolution of the object by the PSF:

$$I = O * PSF \tag{1.25}$$

Fast convolution

The *convolution theorem* states that the Fourier transform of a convolution is the Hadamard product of the Fourier transforms, so that: $\mathscr{F}\{f * g\} = \mathscr{F}\{f\} \cdot \mathscr{F}\{g\}$, where $\cdot$ denotes the Hadamard multiplication. This also works the other way around: $\mathscr{F}\{f \cdot g\} = \mathscr{F}\{f\} * \mathscr{F}\{g\}$. A major application of this theorem is found in the computation of the image as a function of object and PSF:

$$I = O * \mathrm{PSF} = \mathscr{F}^{-1}\{\mathscr{F}\{O\} \cdot \mathscr{F}\{\mathrm{PSF}\}\} \tag{1.26}$$

which tends to be faster than the direct calculation from equation 1.23.

Optical transfer function

The optical transfer function (OTF) of an optical system specifies how different spatial frequencies are transmitted through the system. The optical transfer function is the Fourier transform of the point spread function. When writing the convolution equation in Fourier space, one can see that the spatial frequencies of the image are those of the object filtered by the OTF of the system:

---

[5] In general, a signal can be decomposed as a weighted sum of basis signals. For example, in a Fourier series, any periodic signal, even a rectangular pulse, can be represented by a sum of sine and cosine functions. However, a signal can also be represented as a sum of scaled and shifted impulse (delta) functions. As a demonstration, consider the 1D signal $x(0) = 1, x(1) = 3, x(2) = 2$. Since the impulse function, $\delta(n)$ is 1 at $n = 0$, and 0 at $n \neq 0$, then the signal at $x(1)$ can be written as $3\delta(n-1)$ because $\delta(n-1)$ is 1 at $n = 1$ and 0 elsewhere. Similarly the signals at $x(0)$ and $x(2)$ can be written as $1\delta(n-0)$ and $2\delta(n-2)$). We can thus write the entire signal as a sum of scaled and shifted delta functions, i.e. $x(n) = \sum_k x(k)\delta(n-k)$.

$$\mathscr{F}\{I\} = \mathscr{F}\{O\} \cdot \mathrm{OTF}, \tag{1.27}$$

meaning that the only spatial frequencies from the object that will be kept in the image are the ones which will be non-zero in the OTF. Since the point spread function is the square absolute of the inverse Fourier transformed pupil function, the optical transfer function can also be calculated directly from the pupil function: it is in fact the auto-correlation of the pupil function[6], i.e.,

$$\mathrm{OTF} = P \star P \tag{1.28}$$

where the discrete 2D cross-correlation is given by:

$$(f \star g)[n,m] \ = \sum_p \sum_q f^*[p,q] \; g[p+m, q+n]. \tag{1.29}$$

The OTF is in general a complex function having both a magnitude and a phase portion. The former is referred to as the modulation transfer function (MTF), and the latter the phase transfer function (PTF). The MTF is the magnitude response of the optical system to sinusoids of different spatial frequencies. As such, it provides a useful measure of true or effective resolution, since it accounts for the amount of blur and contrast over a range of spatial frequencies.

Given the shape of the pupil, its MTF can be estimated without computer by doing the autocorrelation by hand as shown on Figure 1.3 for a simple single aperture: simply by sliding the pupil onto of itself, looking for the non-zero intersection, and summing the surface of the intersecting area. The shape of the autocorrelation peaks in the MTF is described by the following equation as a function of the radial frequency u:

$$\mathrm{OTF}(u) = \frac{2}{\pi}\left(\arccos(|u|) - |u|\sqrt{1-u^2}\right). \tag{1.30}$$

Note that frequencies higher than D/λ are absent from the MTF in this case.

Figure 1.4 presents the slightly more complex situation of a pupil with two sub-apertures. In this case the MTF presents two half-strength peaks surrounding the main peak.

1.3 *Sparse aperture characteristics and OTF*

Sparse-aperture pupils are often characterized by their *fill factor*, which is the ratio of the area of a given aperture to the area of a filled aperture

[6] The autocorrelation function and the power spectrum are a Fourier transform pair.

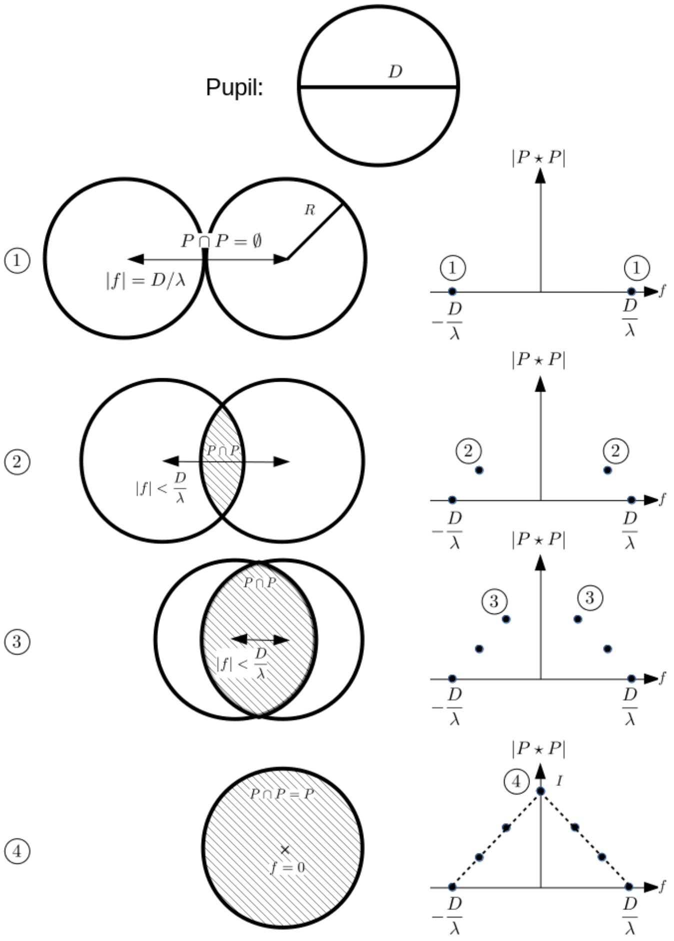
Pupil:
D
R
①
P ∩ P = ∅
|f| = D/λ
②
P ∩ P
|f| < D/λ
③
P ∩ P
|f| < D/λ
④
P ∩ P = P
f = 0
|P ⋆ P|
I
f
−D/λ
D/λ

Figure 1.3: How to build "by hand" the autocorrelation of a single aperture pupil P. For simplicity we assume here the pupil is uniform and real-valued (zero phase). We apply Equation 1.29 by sliding the pupil on top of itself (here from left to right), and computing the integral of the overlapping surface (hashed). Note that since the pupil is point-symmetric, so is the MTF. We obtain the well-known result that the frequency cutoff is $\frac{D}{\lambda}$.

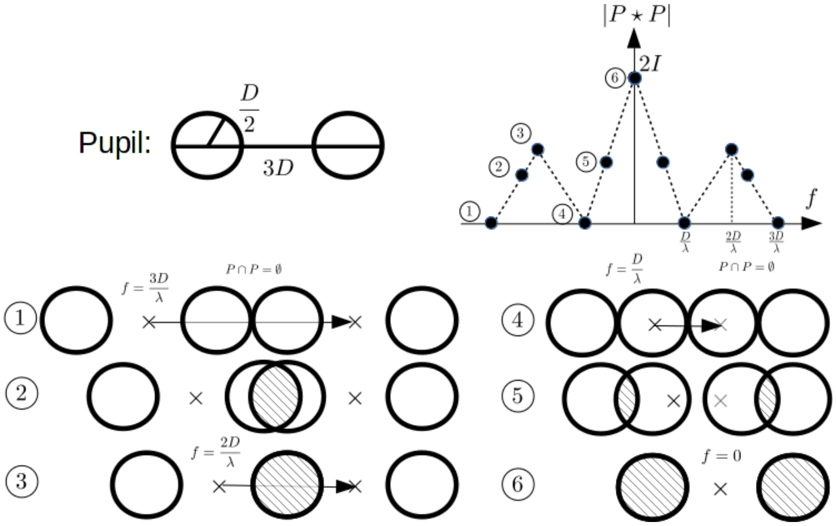
Pupil:
D/2
3D
|P ⋆ P|
2I
f
①
②
③
④
⑤
⑥
D/λ
2D/λ
3D/λ
f = 3D/λ
P ∩ P = ∅
f = D/λ
P ∩ P = ∅
f = 2D/λ
f = 0

Figure 1.4: How to build "by hand" the autocorrelation of a pupil with two sub-apertures. The same principles as for Figure 1.4 apply. This pupil is not point-symmetric, so the shape of the MTF slide depends on the sliding direction: this is obvious since this pupil will provide better resolution along the horizontal axis than the vertical one.

having comparable theoretical resolution. Pupils with low fill factor are said to be *diluted* pupils; conversely, pupils with high fill factor are said to be *compact* pupils.

Non-redundant apertures consists of sub-apertures positioned so that no two pairs of sub-apertures have similar separation vectors, i.e. so that the different pairs do not give rise to overlapping peaks in the OTF. *Redundant* apertures have overlapping frequencies in the OTF.

A desirable feature for sparse aperture imaging is to enhance resolution while minimizing the total light collection area (since large mirrors are expensive). Hence, good non-redundant apertures such as the Golay family of pupils cover the spatial frequency/OTF plane as much as possible without leaving holes (zeros): they are non-redundant yet maximally compact pupils.

1.4 Some practical matters

Convolution and correlation

	Convolution	Correlation
Notation	$g = f * h$	$Corr(f,h) = f \star h$
Fourier Transform pair	$f * h \leftrightarrow FH$ (Convolution Theorem)	$f \star h \leftrightarrow F^* H$ (Correlation Theorem)
Properties	$f * h = h * f$	$f \star h \neq h \star f$
Discrete 1-D description (*)	$g_i = \sum_j f_j h_{i-j}$	$Corr(f,h)_i = \sum_j f_{j-i} h_j$ $(= \sum_j f_j h_{i+j})$
(*) Moving strip approach	h scans f: sequence *reversed*	f scans h: sequence *not* reversed
Example of use	Image formation	Image registration

Table 1.2: Some characteristics of convolution and correlation of two real functions

Autocorrelation

	Autocorrelation
Notation	$Corr(f,f) = f \star f$
Fourier Transform pair	$f \star f \leftrightarrow FF^* = \lvert F \rvert^2$ (Wiener-Khinchin Theorem)
Properties	$f \star f$ is an even function
Discrete 1-D description (*)	$Corr(f,f)_i = \sum_j f_{j-i} f_j$ $(= \sum_j f_j f_{i+j})$
(*) Moving strip approach	f scans f: sequence *not* reversed
Example of use	Generation of OTF from pupil function

Table 1.3: Some characteristics of autocorrelation of a real function

Example: Let $f = \{2\ 2\ 3\ 3\ 4\}$ and $h = \{1\ 1\ 2\}$, then $f * h = \{2\ 4\ 9\ 10\ 13\ 10\ 8\}$, $f \star h = \{4\ 6\ 10\ 11\ 14\ 7\ 4\}$, and $h \star h = \{2\ 3\ 6\ 3\ 2\}$.

Computing convolution and correlation using the discrete Fourier Transform

The convolution and (auto)correlation of two functions can be computed using the Fourier Transform and the Fourier Transform pair relations listed in tables 1.2 and 1.3. However, it is important to ensure that the array sizes used in the transforms are large enough to accommodate the resultant convolution/correlation which has a size in each dimension that is equal to the sum of the maximum sizes of each of the two functions in the same dimension. Practically, this is achieved by padding the arrays that describe the two functions to be convolved/correlated with zeros.

Image sampling

A digital image represents a sampled version of a continuous optical image. Converting from a continuous optical image $g(x,y)$ to its digital representation $g_{obs}(x_i,y_i)$ requires averaging over the pixel area and sampling. That is,

$$g_{obs}(x_i,y_i) = \{(x,y) * rect(\Delta x',\Delta y')\} \times \delta(x,y : \Delta x,\Delta y) \tag{1.31}$$

where $rect(\Delta x',\Delta y')$ and $\delta(x,y : \Delta x,\Delta y)$ are the rect and sampling functions, respectively, $\Delta x, \Delta y$ are the sampling intervals (i.e. the spacing between the pixel centers in the x- and y-directions), and $\Delta x' \leq \Delta x, \Delta y' \leq \Delta y$ are the dimensions of the photosensitive region of a pixel. Note, the fraction of the detector surface that is sensitive to light is called the fill factor (unrelated to the pupil's fill factor !) and is given by:

$$\textit{fill factor} = 100\% \times (\Delta x'\Delta y')/(\Delta x\Delta y). \tag{1.32}$$

According to the Nyquist-Shannon sampling theorem, in order to preserve the spatial resolution of the original continuous image, the sampling intervals should be no greater than one-half the size of the smallest resolvable feature of the optical image. This "critical" (or "Nyquist"), sampling is equivalent to acquiring samples at twice the highest spatial frequency contained in the image. We note that the use of a smaller sampling interval does not yield more spatial information. However, this "over-sampling" is relatively harmless, the only downsides being increased storage requirements and computation times for image processing. The same is not true when using a larger sampling interval. Here the subsequent 'under-sampling' can be highly detrimental as it results in both a loss of information and aliasing artifacts in the recorded imagery. Aliasing is where high spatial frequency signals appear as low spatial frequency signals in the sampled signal.

For critical sampling of an image acquired with an instrument with a circular aperture we have $2\Delta x = 1.22\lambda/D = 2\Delta y$. The Fourier

spectrum of the sampled image is a band-limited function, due to the multiplication of the object spectrum by the OTF, with cut-off spatial frequencies of $u_{co} = 1/(2\Delta x)$ and $v_{co} = 1/(2\Delta y)$, and frequency resolutions of $\Delta u = 1/(N^x_{pix}\Delta x)$ and $\Delta v = 1/(N^y_{pix}\Delta y)$, where N^x_{pix}, N^y_{pix} are the number of pixels in the camera along the x- and y- directions, respectively.
As an example, if a detector with an array of $N_{pix} \times N_{pix}$ square pixels provides critical sampling of an image acquired through a circular aperture at wavelength λ_1, then the diameter of the OTF, on the same pixel grid, is N_{pix} and the diameter of pupil is $N_{pix}/2$. To simulate the PSF for an image acquired with the same imaging system and detector at another wavelength $\lambda_2 > \lambda_1$ we have to scale the size of the OTF, and thus the pupil, by (λ_1/λ_2) because the cut-off frequency varies with wavelength ($\propto D/\lambda$). Note, for $\lambda_2 < \lambda_1$ the image would be under-sampled.

EXERCISE 1.3—

1. Convolve the two 1D functions $(3, 4, 5)$ and $(2, 1, 3)$ using the serial product and the FFT methods.
2. Generate a PSF via the autocorrelation of the pupil function.
3. Show how the images of M31, Saturn, and the USAF 1951 pattern are affected by having 1 wave of defocus in the imaging system.
4. Show how an image of a triple star system is affected by having $\frac{1}{2}$ wave of vertical coma in the imaging system.
5. Compare the MTFs for 1 wave of defocus and $\frac{1}{2}$ wave of vertical coma.
6. Show that the Golay-6 pupil is a good pupil for imaging. How can we check this on its OTF ?
7. Make an image of an object through the Golay-3 pupil. What happens when we increase the sub-aperture distance ?
8. How can we determine the size of the aperture used to generate an image, from just the image?

1.5 *Atmosphere and turbulence*

The wavefront from an astronomical source can be considered flat at the top of the atmosphere. But as it propagates to the ground the optical turbulence corrupts its phase and amplitude, affecting measurements from ground-based telescopes. The perturbations of the optical path length $l(\vec{x})$ are considered as achromatic[7]. However, the phase of the optical wave $\phi(\vec{x}) = \frac{2\pi}{\lambda} l(\vec{x})$ strongly depends on the wavelength λ.

Optical turbulence is caused by the mechanical mixing of layers of air with different temperatures and hence density. As the refractive index of air changes with density this turbulence creates a continuous screen of spatially and temporally varying refractive indices. Although each of the refractive index inhomogeneities in the turbulent layers may be small the wavefront passes through a large number of them and the cumulative effect can be quite large. The cumulative refractive index variations delay parts of the incoming wavefront with respect to others. The net effect is that the wavefront becomes aberrated. If we assume a horizontal turbulent layer at altitude, h, above the ground and that the layer thickness, δh, is large compared to the eddy size of the refractive index inhomogeneities, but small enough so that we can ignore diffraction effects within the layer (thin screen approximation), then the phase fluctuations, $\phi(r)$, induced by the turbulent layer is related to the refractive index fluctuations, $n(h,r)$, along the propagation path by

---

[7] Air is slightly dispersive, but this is usually neglected.

$$\phi(r) = k \int\limits_{h}^{h+\delta h} n(h,r)dh \tag{1.33}$$

where k is the wave number, $2\pi/\lambda$, with λ being the wavelength of the light and r is a spatial coordinate. The wave function after the layer is then,

$$\psi(r) = \chi(r) \exp i\phi(r) \propto \exp i\phi(r) \tag{1.34}$$

It is these aberrations in the wavefront which act to distort images from astronomical telescopes. Now the refractive index fluctuations are random and the spatial structure of a random process is given by *structure functions*. The structure function $D_f(r_1, r_2)$ of a random variable f measured at positions $r_1,\ r_2$ is defined by

$$D_x(r_1, r_2) = \left\langle \left| f\left(r_1\right) - f\left(r_2\right) \right|^2 \right\rangle, \tag{1.35}$$

where the angled brackets, $\langle ... \rangle$, represent the spatial average. That is, the structure function measures the expectation value of the difference of the values of f measured at two positions, r_1 and r_2. With this in mind, we are therefore not interested in the absolute value of the phase, only the difference between its value at two points, which is caused by the spatial variance of the refractive index. The refractive index structure function, $D_n(\rho)$, is thus the spatial variance in the difference of refractive index n as a function of separation ρ.

$$D_n(\rho) = \left\langle \left| n\left(r\right) - n\left(r+\rho\right) \right|^2 \right\rangle. \tag{1.36}$$

The Kolmogorov model of turbulence

A description of the nature of the wavefront perturbations introduced by the atmosphere is provided by the Kolmogorov model developed by Tatarski, based partly on the studies of turbulence by the Russian mathematician Andrei Kolmogorov. This model is supported by a variety of experimental measurements and is widely used in simulations of astronomical imaging. The model assumes that the wavefront perturbations are brought about by variations in the refractive index of the atmosphere. In the Kolmogorov model of turbulence the refractive index structure function is given by

$$D_n(\rho) = C_n^2\, \rho^{2/3} \quad \text{for } \; l_0 < \rho < L_0 \tag{1.37}$$

where the quantity C_n^2 is known as the refractive index structure coefficient (see section below on C_n^2) and can be considered to be a measure of the small-scale turbulence, and l_0 and L_0 are the inner and outer scales of turbulence, respectively. The former is the size below which viscous effects and energy is dissipated into heat. The latter is the size above which isotropic behavior is violated. We note that the inner scale size of turbulence l_0 varies inversely with C_n^2 and has values that range in size from 1 - 2 mm up to 1 cm or more near the ground.

The power spectral density for refractive index fluctuations over inertial subrange ($1/L_0 < \kappa < 1/l_0$) in three dimensions is defined by

$$P_n(\kappa) = 0.033\, C_n^2\, \kappa^{-11/3} \tag{1.38}$$

where κ is the spatial frequency.

The refractive index variations lead directly to phase fluctuations described by $\phi\,(\mathbf{r})$, but any amplitude fluctuations are only brought about as a second-order effect while the perturbed wavefronts propagate from the perturbing atmospheric layer to the telescope. For all reasonable models of the Earth's atmosphere at optical and infra-red wavelengths the instantaneous *imaging* performance is dominated by the phase fluctuations $\phi\,(\mathbf{r})$. The amplitude fluctuations described by $\chi\,(\mathbf{r})$ have negligible effect on the structure of the images seen in the focus of a large telescope, but will of course affect *photometric* measurements.

As the phase perturbations are random, we use the phase structure function to describe the spatial structure of the phases, i.e.

$$D_\phi\,(\rho) = \left\langle \left|\phi\,(\mathbf{r}) - \phi\,(\mathbf{r} + \rho)\right|^2 \right\rangle_{\mathbf{r}} \tag{1.39}$$

where $D_\phi\,(\rho)$ is the atmospherically induced variance between the phase at two parts of the wavefront separated by a distance ρ in the aperture plane, and $\langle \ldots \rangle$ represents the ensemble average.

The power spectral density for phase fluctuations is given by[8]

$$P_\phi(\kappa) = |\mathscr{F}\,(\phi)\,|_\kappa^2 = 0.023\, r_0^{5/3}\, \kappa^{-11/3}. \tag{1.40}$$

---

[8] Modeling atmospheric distortion is done in three steps. First we create a 2D complex array of random numbers having a Gaussian distribution with mean = 0 and sigma = 1. Second, we multiply this array by the modulus (amplitude) of the wavefront phase spectral density. Third, we Fourier Transform the resultant array: the real component is a realization of a Kolmogorov phase screen. IMAGE: SIMULATED TURBULENCE PHASE SCREEN

C_n^2 *profile*

A more thorough description of the astronomical seeing at an observatory is given by producing a profile of the turbulence strength as a function of altitude h, called a C_n^2 profile.

The model most often used is the generalized Hufnagel-Valley (HV) model:

$$\begin{aligned} C_n^2(h) &= A\exp\left(-h/H_A\right) + B\exp\left(-h/H_B\right) \\ &+ C\,h^{10}\exp\left(-h/H_C\right) \\ &+ D\exp\left(-(h-H_D)^2/2d^2\right) \end{aligned} \tag{1.41}$$

Here A is the coefficient for the surface (boundary layer) turbulence strength ($\mathrm{m}^{-2/3}$) and H_A is the height for its *1/e* decay (meters), B and H_B similarly define the turbulence in the troposphere (up to about 10 km), C and H_C define the turbulence peak at the troposphere, and D and H_D define one or more isolated layers of turbulence, with d being the layer thickness (meters). The coefficients for the HV(5-7) model are[9]: $A = 1.7 \times 10^{-14}\ \mathrm{m}^{-2/3}$, $H_A = 100;\ \mathrm{m}$, $B = 2.7 \times 10^{-16}\ \mathrm{m}^{-2/3}$, $H_B = 1500\ \mathrm{m}$, $C = 3.59 \times 10^{-53}\ \mathrm{m}^{-2/3}$; $H_C = 1000\ \mathrm{m}^{-2/3}$, $D = 0\ \mathrm{m}$. This model highlights the fact that the dominant locations for index of refraction fluctuations that affect astronomers are the atmospheric boundary layer and the troposphere. However, each observing site will have its own set of typical values for the coefficients.

Values of C_n^2 of $\sim 10^{-17}\mathrm{m}^{-2/3}$ or less indicate weak turbulence, while values of $\sim 10^{-13}\mathrm{m}^{-2/3}$ or more indicate strong turbulence.

C_n^2 profiles are generally measured when deciding on the type of adaptive optics system which will be needed at a particular telescope, or in deciding whether or not a particular location would be a good site for setting up a new astronomical observatory. Typically, several methods are used simultaneously for measuring the C_n^2 profile and then compared.

[10]

Fried parameter: r_0

For simplicity, the phase fluctuations in Tatarski's model are often assumed to have a Gaussian random distribution. This is because for atmospheric layers that are much thicker than the turbulence cells, many independent variables contribute to the phase shift (see Equation 1.33). Therefore the Central Limit Theorem implies that ϕ has Gaussian statistics. In this case the structure function can be described in terms of a single parameter r_0, i.e.,

$$D_\phi(\rho) = 6.88\left(\frac{|\rho|}{r_0}\right)^{5/3}. \tag{1.42}$$

[9] Hardy, 1998

[10] IMAGE: C_N^2 PROFILE

The parameter r_0, which is often known as the Fried parameter, named after David L. Fried, indicates the strength of the phase fluctuations and is frequently used to describe the atmospheric conditions at astronomical observatories. Although often regarded as a constant, r_0 is a statistical parameter that varies considerably over short periods of time, sometimes changing by a factor of 2 within a few seconds. [11] r_0 can be determined from a measured C_n^2 profile using:

$$r_0 = \left[0.423\, k^2 (\cos z)^{-1} \int_0^\infty C_n^2(h) dh\right]^{-3/5} \tag{1.43}$$

where the turbulence strength $C_n^2(h)$ varies as a function of height h above the telescope, and z is the angular distance of the astronomical source from the zenith. The key points of equation (1.43) are that the optical effects of turbulence are less (i.e. r_0 is larger) at smaller zenith angles and at longer wavelengths. Equation (1.43) shows that r_0 varies with wavelength as $\lambda^{6/5}$. We note the standard wavelength for reporting r_0 values is 0.5 μm.

A commonly used definition for r_0 is that it corresponds to the aperture diameter D for which the variance σ^2 of the wavefront phase averaged over the aperture comes approximately to unity:

$$\sigma^2 = 1.0299 \left(\frac{D}{r_0}\right)^{5/3} \tag{1.44}$$

Equation (1.44) shows that r_0 corresponds to the length-scale over which the phase perturbations due to atmospheric turbulence becomes significant (15 to 20 cm at visible wavelengths at good observatories).

We note that r_0 is often referred to as the characteristic scale of atmospheric turbulence. This is not quite true as the Kolmogorov law does not have any characteristic scale. However, only the perturbations of order r_0 are relevant for long-exposure imaging (see section 1.6).

Atmospheric time constant: τ_0

A second parameter that is used to describe the turbulence conditions at a site is the atmospheric time constant τ_0. This parameter, which is also know as the atmospheric coherence time and the Greenwood time constant (after Darryl Greenwood), corresponds to the time-scale over which the the changes in the turbulence become significant. This occurs when the wavefront phase decorrelates by ~ 1 radian RMS). If the evolution of the atmosphere is slow enough that it can be considered "frozen" at short time scales, and the turbulence is just advected across the aperture by the wind in the atmosphere (the Taylor Hypothesis), then τ_0 is simply proportional to r_0 divided by the mean wind speed, i.e.

$$\tau_0 = 0.31 r_0 / \bar{V} \tag{1.45}$$

[11] IMAGE: Graph of r_0 verses time showing variability on short time scales, as well as r_0 spread during an observing night

where

$$\bar{V} = \left[\frac{\int C_n^2(h)|V(h)|^{5/3} dh}{\int C_n^2(h) dh} \right]^{3/5} . \tag{1.46}$$

Equation (1.46) shows that the main contribution to $\bar{V}$ is from high-velocity winds where C_n^2 is large.

Images taken with exposure times $\leq \tau_0$ are called "short-exposure" images and their phase aberrations are considered frozen (fixed). At longer exposure times the aberrations are averaged and for exposure times $>> \tau_0$ we obtain the long-exposure PSF. Like r_0, τ_0[12] varies with wavelength as $\lambda^{6/5}$.

Isoplanatic angle: θ_0

The viewing angle over which the wave front error can be considered constant is called the *isoplanatic* angle.

However, atmospheric turbulence is continuously changing and the instantaneous phase aberrations observed in the aperture therefore change with the viewing direction. That is, when beams propagate in different directions through a series of turbulence layers, each beam will sample slightly different turbulence. Propagation through turbulence is therefore, in general, *anisoplanatic*.

The isoplanatic angle is given by

$$\theta_0 = \left[2.91 k^2 \int_{path}^{h_L} C_n^2(h) h^{5/3} dz \right]^{-3/5} , \tag{1.47}$$

where h_L is the height of the turbulent layer. For constant C_n^2 equation(1.47) reduces to

$$\theta_0 \simeq 0.6 \, r_0 / h_L . \tag{1.48}$$

We note that the isoplanatic angle is hugely variable and in extreme turbulence situations can approach the diffraction limit of the telescope.[13] Like r_0 and τ_0, θ_0 varies with wavelength as $\lambda^{6/5}$.

Layer model for atmosphere

For simplicity we commonly approximate the C_N^2 profile by a number of layers where each layer propagates in the direction of the local wind velocity and is characterized by a turbulence strength that is constant within the layer. In addition, the thickness of a layer is assumed to be large compared to the scale size of the turbulence cells, so that Gaussian statistics apply, but still small enough for diffraction effects to be negligible.

[12] r_0 determines the spacing of the actuators needed in an adaptive optics system, and τ_0 determines the correction speed required to compensate for the effects of the atmosphere.

[13] For angular separations smaller than the isoplanatic angle, the variance between difference paths is small and we can treat the PSF as spatially invariant and model the image as a convolution of the object and PSF. However, for angular separations larger than the anisoplanatic model, the PSF varies spatially across the field-of-view (FOV) and the convolution model is only approximate, with the approximation becoming worse as the FOV increases.

The Fried parameter for an individual turbulence layer r_{0_i} is related to the overall Fried parameter r_0 by

$$r_0^{-5/3} = \sum_{i=1}^{N} r_{0_i}^{-5/3}. \qquad (1.49)$$

Scintillation

When the propagation path length L from a turbulent layer is comparable to, or longer than, the Fresnel propagation length

$$L_F = r_0^2/\lambda\ , \qquad (1.50)$$

the rays diffracted at the turbulence cells in the layer interfere with each other which causes intensity variations, or *scintillation*, in the pupil plane in addition to phase variations. For turbulence conditions of r_0 = 20 cm and $\lambda = 0.5\mu$m, the Fresnel propagation length L_F = 80 km. This is significantly larger than the height of the layers that provide the major contributions to the C_N^2 integrals, and scintillation effects are negligible on imaging compared to phase effects as long as the zenith angle is not too large.

The characteristic size of the scintillation speckles is given by the Fresnel scale:

$$r_F = \sqrt{L\lambda}. \qquad (1.51)$$

As a wavefront propagates away from a turbulent layer, increasing L and hence r_F, the spatial intensity fluctuations become larger in terms of spatial extent. This is not dependent on the strength of the layer, which only affects the magnitude of the intensity fluctuations and not their spatial properties; for example, a turbulent layer at 10 km observed at 500 nm results in speckles of size $\sim$ 7 cm, irrespective of the strength of the turbulence.

Equation (1.50) is equivalent to saying when the Fresnel scale r_F is commensurate with or larger than the Fried scale r_0 we have to consider the effects of scintillation. As $L = h_L/\cos z$, where h_L is the height of the turbulent layer, it is primarily the high altitude turbulent layers in the atmosphere and observations at high zenith angles that contribute to the intensity variations of the wave front. The resulting 'scintillation noise' is what limits the precision of photometric measurements.

The effects of this noise can be quantified by determining the variance of intensity fluctuations $\sigma_I^2 = \frac{\langle I^2\rangle - \langle I\rangle^2}{\langle I\rangle^2}$. As the scintillation noise is caused by the intensity speckles entering and leaving the telescope pupil there are three regimes of scintillation noise: short exposures and longer exposures on medium/large telescopes, and the specific case of small telescopes.

- For short exposures the intensity speckles appear frozen in the pupil and no temporal averaging occurs. The expression of the variance is shown to be [14]:

  $$\sigma_I^2 = 17.34\, d^{-7/3} \left(\cos z\right)^{-3} \int C_n^2(h) h^2 dh \,. \tag{1.52}$$

  where d is the diameter of the telescope. Note how this expression shows the expected strong decrease of scintillation amplitude with aperture size and that it is independent of the observing wavelength.

- For longer exposures the speckles traverse the pupil during an exposure and the scintillation noise will be reduced by temporal averaging. The amount by which the scintillation noise is reduced is dependent on the wind speed:

  $$\sigma_I^2 = 10.66\, d^{-4/3} t^{-1} \left(\cos z\right)^{\alpha} \int \frac{C_N^2(h) h^2}{V_\perp(h)} dh \,, \tag{1.53}$$

  where $V_\perp(h)$ is the wind velocity profile and α the airmass exponent (taken to be ~ -3.5 in practice). Here again scintillation amplitude decreases with aperture size, and no wavelength dependency is apparent.

- For telescopes smaller than a few tens of centimeters ($D < r_F$) there will be a significant wavelength dependence since the size of the speckles is itself wavelength-dependent.[15]

  $$\sigma_I^2 = 19.2\, \lambda^{-7/6} \left(\cos z\right)^{-11/6} \int C_N^2(h) h^{5/6} dh \,, \tag{1.54}$$

Another way to look at scintillation is as a fluctuation in the spatial intensity distribution in the pupil caused by the random turbulent atmosphere locally focusing and defocussing the wave front. Scintillation is therefore a second order effect as it depends on the curvature (i.e., second derivative) of the wave front, whereas phase perturbations in the pupil plane are a first-order effect as they depend on the first derivative of the wave front.

Finally, Young's approximation is a semi-empirical formula that gives an estimation of the scintillation noise variance:

$$\sigma^2 = 10^{-5} C^2 d^{-\frac{4}{3}} t^{-1} \left(\cos z\right)^{-3} \exp\left(-2h_{obs}/H\right) \tag{1.55}$$

where d is the diameter of the telescope, t is the exposure time of the observation, z is the zenith distance, h_{obs} is the altitude of the observatory and H the scale height of the atmospheric turbulence, which is generally accepted to be approximately 8000 m. C has been empirically measured at several major observatories to have a mean value of $\simeq 1.5$ in SI units [16].

[14] Kenyon et al., 2006

[15] Dravins et al., 1998

[16] Osborn et al., 2015

1.6 *Imaging through atmospheric turbulence*

The atmosphere has two distinct states of motion: laminar and turbulent. Mixing of the air does not occur in laminar flow but turbulent flow is characterized by dynamic mixing and it acquires random subflows called eddies. The different size eddies cause different imaging effects. Eddies larger than the beam size cause tilt variations (beam wander) while eddies smaller than the beam size cause it to break up as different sections of the beam are refracted in different directions (beam spread). That is, the atmosphere can be visualized as a path of continually changing lenses refracting light as it travels to the telescope.[17] In other words, the atmospheric turbulence may be regarded as random phase aberrations added to the telescope. These aberrations constantly change in time, and consequently so does the PSF.

The wave front variance for beam spread is found by removing the contribution from the two-axis tilt variations from equation (1.44) and is given by

$$\sigma^2_{\mathrm{tiltcomp}} = 0.134 \left(\frac{D}{r_0}\right)^{5/3} \tag{1.56}$$

This is the origin of the common statement that "tip-tilt is the largest contributor to wave front error". However, if (D/r_0) is large, removing tip/tilt won't yield anywhere near a diffraction-limited image.

Equations (1.44) and (1.56) show that the ratio D/r_0 provides a measure of the strength of the atmospheric turbulence. When the quantity D/r_0 is a few we have low turbulence. When D/r_0 is large we have strong turbulence.

Let's now take a look at the morphology of a short-exposure, turbulence degraded, image. If we consider r_0 as being roughly the average physical dimension of a single isoplanatic "patch", and a patch gives rise to a diffraction-limited "speckle" in the PSF of width $2.44\lambda/D$, then the total number of speckles in the short-exposure PSF from all the patches in the telescope pupil of diameter D is $\sim (D/r_0)^2$ and they are spread over an area of diameter $\sim 2.44\lambda/r_0$. The latter is known as the "seeing" disk and its location depends on the tip/tilt perturbations in the wave front. In summary, there are two components that contribute to the morphology of a short-exposure turbulence-degraded image: the diffraction limit of the aperture which depends on λ/D, and the short-exposure image spread which depends on λ/r_0. For long-exposure images there is a third component related to the image displacement caused by non-zero tip/tilt aberrations in the atmospheric wave front. This component depends on both D and r_0.

[18]

[17] This is only illustrative: there are no discontinuities in the atmosphere.

[18] IMAGE: TWO ROWS WITH INSTANTANEOUS "SPECKLE" PSFS IN THE TOP ROW AND THE LONG-EXPOSURE (500 IMAGES) PSFS IN THE BOTTOM ROW. FOUR COLUMNS FOR R0 OF 50cm, 20cm, 10cm, AND 5cm WITH A TELESCOPE OF APERTURE D=100 cm

Long-exposure images

Long-exposure images are images that are acquired over exposure times that are long compared to the atmospheric coherence time t_0.

The diameter of a long-exposure PSF is the sum of the three contributions described at the end of above section. The long-exposure PSF, which has the three components described at the end of the last section, can modeled using two functions representing a "core" and a "halo". We'll first take a look at the "core" component.

The mean-square deflection angle, or angle of arrival, due to tip/tilt variations has a Gaussian distribution with a variance given by

$$\alpha^2 = 0.18 \left(\frac{\lambda}{D}\right)^2 \left(\frac{D}{r_0}\right)^{5/3} . \tag{1.57}$$

That is, the mean-square deflection angle varies as $\sim D^{-1/3}$ and is independent of λ. However, relative to the Airy disk, image motion gets worse for larger D and smaller λ, i.e.

$$\frac{\alpha}{\lambda/D} = 0.43 \left(\frac{D}{r_0}\right)^{5/6} \propto \frac{D^{5/6}}{\lambda}. \tag{1.58}$$

The effect of a random tilt error α is to spread each point of an image into a Gaussian profile with standard deviation α. Therefore the FWHM of the central core of a long-exposure PSF, including the tilt error, can be approximated by

$$\begin{aligned} FWHM_{core} &= 2.354 \sqrt{\alpha^2 + 0.191\,(\lambda/D)^2} \\ &= 1.03 \left(\frac{\lambda}{D}\right) \sqrt{1 + 0.94\,(D/r_0)^{5/3}} \end{aligned} \tag{1.59}$$

where the second term represents the variance of a Gaussian profile with the same FWHM as an Airy function (see section 1.1).

Equation (1.59) shows that for $D/r_0 \sim 1$ the core of the PSF is close to diffraction-limited and almost all the long-exposure degradation to the image is caused by image motion (wander). As the ratio D/r_0 increases, both the image motion and the number of speckles increase and the FWHM of the core of the long-exposure PSF approaches that of the halo from the short-exposure image spread, i.e. $0.98\lambda/r_0$.[19]

Thus, for telescopes with diameters $D \sim r_0$ or smaller, the resolution of long-exposure images is determined by diffraction and is given by $\sim 1.22\lambda/D$. For telescopes with D greater than r_0, the long-exposure image resolution is given by $\sim 1.22\lambda/r_0$. That is, the resolution is determined primarily by the atmosphere and is independent of telescope diameter, remaining constant at the value given by a telescope of diameter equal to r_0.[20] We note that the long-exposure PSF is independent of the viewing direction (i.e. it is isoplanatic) because the turbulence

[19] The FWHM of the long-exposure PSF is known in ground-based astronomy as the "seeing" parameter $\beta(\lambda) = 0.98\lambda/r_0$. In arcseconds, $\beta(\lambda) \simeq 20\lambda/r_0$. where λ is in microns and r_0 is in centimeters.

[20] Due to the wavelength dependence of r_0, the "seeing" depends on $\lambda^{-1/5}$ and we can obtain slightly higher-resolution long-exposure imagery by observing at longer wavelengths when using large telescopes.

and its structure function are statistically the same everywhere in the field.

Atmospheric transfer function

For long exposure images, theory leads to the expression.

$$\mathrm{OTF}_{\mathrm{LongExposure}} = \mathrm{OTF}_{\mathrm{Tel}} \cdot \mathrm{OTF}_{\mathrm{Atm}} \tag{1.60}$$

For large telescopes with good optical quality the resolution is entirely defined by the atmosphere, so we neglect the first term and $OTF_{\mathrm{LE}} \approx OTF_{\mathrm{Atm}}$. The atmospheric OTF is related to the statistics of atmospheric phase aberrations through the phase structure function $D_{\phi}(\vec{r})$:

$$OTF_a(\vec{f}) = \exp[-0.5 D_{\phi}(\lambda \vec{f})]. \tag{1.61}$$

Speckle imaging

As discussed at the beginning of this section, the instantaneous atmospheric PSF consists of a number of "speckles" that depends on the level of turbulence as measured by the ratio D/r_0. Thus, imaging at frame rates that are commensurate with, or less than, the atmospheric coherence time is referred to as *speckle imaging*. The main advantage of speckle imaging is that each frame contains spatial information out to the diffraction limit of the telescope. Speckle imaging can be divided into two approaches: "shift and add" (imaging stacking) and speckle interferometry.

Shift-and-add

One approach to recovering the high spatial-frequency information in a set of speckle images is to shift all the short-exposure images to a common center-of-mass and then average them. This procedure is effectively removing the image motion due to the tip/tilt components of the atmospheric wave front. However, it has been found that a better approach which provides a higher Strehl ratio, at least for images of objects with point-like structures, is to align the images to the brightest speckle and then average.

For low levels of turbulence the best resolution is provided by a variation of the shift-and-add approach, known as *"lucky imaging"*. Since atmospheric turbulence is a random process there are moments in time when the seeing is significantly better or worse than average. If we only select the speckle images where the seeing is essentially diffraction-limited ("lucky images", we can produce a better image than would be possible by combining all the images without frame selection.

The probability that we can observe an image with diffraction-limited resolution is given by:

$$\mathrm{P_{Lucky}} \simeq 5.6\ \exp\left[-0.1557\,(D/r_0)^2\right] \ \text{for} \ \ D/r_0 \geq 3.5 \tag{1.62}$$

This shows that for turbulence with $D/r_0 = 6$ there is a 1 in 50 chance of acquiring an image with diffraction-limited resolution. For higher levels of turbulence, the probability diminishes quickly: at $D/r_0 = 10$ the probability is reduced to ~1 in 1,000,000. That is, lucky imaging is a technique that only feasible for relatively benign turbulence conditions with $D/r_0 \leq 7$.
21

Speckle Interferometry

Speckle interferometry yields the high resolution autocorrelation of the object and is a popular technique for studying binary stars.

Since $I = O * \mathrm{PSF}$ and the PSF irradiance varies randomly while the object irradiance remains constant, then

$$\left\langle |\mathscr{F}\{I\}|^2 \right\rangle = \left\langle |\mathscr{F}\{O \cdot PSF\}|^2 \right\rangle = |\mathscr{F}\{O\}|^2 \cdot \left\langle |\{\mathscr{F}\{PSF\}|^2 \right\rangle \tag{1.63}$$

where $< .. >$ denotes the ensemble average. Thus the autocorrelation of the object is computed from the recorded speckle data in two steps. First, we average the modulus square of the Fourier transforms (i.e. the power spectra) of all the speckle images and then compensate the resulting spectrum for the transfer function of the atmosphere and instrument. The latter can come from either a model or (better) from a time series of speckle images of an unresolvable star close to the object. This step yields the power spectrum of the object, what Michelson observed as "visibility".

Second, we take the inverse Fourier transform of the corrected ensemble averaged power spectrum to provide the autocorrelation function of the object with a resolution limited only by diffraction, not anymore by the turbulent atmosphere. That is,

$$\mathrm{ACF} = \mathscr{F}^{-1}\left\{\left\langle |\mathscr{F}\{I\}|^2\right\rangle / \left\langle |\mathscr{F}\{PSF\}|^2\right\rangle\right\} = \mathscr{F}^{-1}\left\{|\mathscr{F}\{O\}|^2\right\} \tag{1.64}$$

Note that the ratio in equation (1.64) can only be taken in regions of frequency space where $\left\langle |\mathscr{F}\{PSF\}|^2\right\rangle$ is non-zero and has signal-to-noise greater than unity. The quantity $\left\langle |\mathscr{F}\{PSF\}|^2\right\rangle$ is referred to as the speckle transfer function. The speckle transfer function relates the modulus squared of the object spectrum to the second moment of the image spectrum. The speckle transfer function is finite out to spatial

[21] IMAGE: AN EXAMPLE RAW IMAGE OF A BINARY STAR ACQUIRED THROUGH $D/R_0 = 5$ TURBULENCE, THE AVERAGE OF 500 IMAGES TURBULENCE, THE SAA IMAGE BASED ON CENTER-OF-MASS, THE SAA IMAGE BASED ON BRIGHTEST SPECKLE, AND THE SAA IMAGE BASED ON THE TOP 5% LUCKY IMAGES.

frequencies approaching the diffraction-limited cut-off frequency, and hence allows diffraction-limited information to be measured.

[22]

If we want to reconstruct an actual image of the object from the recorded speckle images, then we can do so if the object and the point source are in the same isoplanatic patch, and they are imaged at the same time (i.e. simultaneously). The approach, called *speckle holography*, uses the images of the point source as keys for the deconvolution of the images of the object. Deconvolution is discussed in the next chapter. We can also reconstruct an image of the object by using *speckle masking* to recover the spectral phase information on the object from the speckle data. Speckle masking is a method which involves estimation of the closure phases from each of the short exposure speckle images. This will be discussed in the chapter on interferometry. The object spectral phase information is combined with its spectral amplitude information from speckle interferometry and inverse transformed to yield the image of the object.

[22] IMAGE: AVERAGE POWER SPECTRUM OF DATA USED IN SAA EXAMPLE, ACF OF DATA, ACF OF RAW DATA CORRECTED BY TRANSFER FUNCTION

2

Image Restoration

Achieving the highest possible resolution imagery from a ground-based telescope requires removing the deleterious effects of image blur. This can be accomplished in two ways: adaptive optics compensation of the incoming wave front (see chapter 3) and numerical image restoration. Ideally both approaches are used in combination. However, as numerical restoration is more accessible, in that it only requires a computer and not an assembly of complex optical hardware, we will address it first.

Image preprocessing

Before we embark on a journey into methods for image restoration we need to point out the success of any type of image restoration depends *strongly* on the quality of the image preprocessing performed beforehand. The types of preprocessing can be divided into three categories: calibration, windowing, edge extension and sectioning.

1. Calibration: Raw image data are subject to variations in the pixel-to-pixel sensitivity of the detector and/or by distortions in the optical path and sit on a pedestal due to the dark current signal in the detector. These effects result in artifacts in the image need to be taken into account, either before embarking in image restoration, or as a part of the image restoration process. The correction for pixel-to-pixel variations is known as a "flat-field" correction and requires measurement of a source of light with constant, uniform intensity across the telescope pupil. The dark current signal is obtained by acquiring images with no incident light. We note that dark current can vary with time if the temperature of the detector changes. The raw data are processed by subtracting a suitable dark frame and dividing by the flat-field frame (which has been dark current corrected).

2. Windowing and edge extension: There is a tendency for deconvolu-

tion to be complicated by the truncation of the blurred image by the recording frame. For methods that require the Fourier transform of the data, we need to take into account that the Fourier Transform always treats an image as if it were part of a periodically replicated array of identical images extending horizontally and vertically to infinity. Therefore when the recording frame truncates the blurred image signal, we need to ensure that this periodicity requirement is met if we are to minimize artifacts due to 'edge effects'. The simplest approach is to "window" (or "apodize") the data frame by multiplying by a function that smoothly goes to zero at the edge of the frame. There are a large number of window functions that satisfy this condition, each with advantages and disadvantages. The Tukey window, also known as the tapered cosine window, is a common choice. The Tukey window function can be regarded as a unit amplitude cosine lobe of width $\alpha N/2$ that is convolved with a unit height rectangular window of width $(1\alpha/2)N$. That is, the data near the edge of the frame are modified by a cosine roll-off such that the signal is zero at the edges of the data frame. Thus, when using a Tukey window to apodize the data we need to bear in mind that the restored data will be corrupted over a distance equivalent to the sum of the cosine lobe width and half of the extent of the PSF from the edge of the frame.

Another approach, which allows for a slightly larger region of the data frame to be accurately restored, is to use edge extension. Simple edge extension involves embedding the data frame in a larger array and then continuing the signal at the edge of the data frame out along straight lines perpendicular to the edge of the frame out to the edge of the larger array and ensuring that the extended signal smoothly tapers to zero at the edge of the larger array. In this case, only the region within a half of the width of the PSF at the edges of the restored image is corrupted.

3. Sectioning: This involves dividing an image with a spatially varying PSF into sections each of which may be treated as having an approximately spatially invariant PSF.

Multiplicative deconvolution

In chapter 1 we saw that images are described by a convolution of an object intensity distribution with a PSF, i.e.

$$I = O * PSF \,, \tag{2.1}$$

which in the Fourier domain can be expressed as

$$\mathscr{F}\{I\} = \mathscr{F}\{O\} \cdot OTF \,. \tag{2.2}$$

To obtain the original object from the blurred image requires knowing the blurring function and performing deconvolution. Given the PSF and the image I it seems trivial to solve for the object O: as long as there are no zeros in the OTF we just use the simple inverse filter $W = 1/OTF$, i.e.,

$$O = \mathscr{F}^{-1}\{\mathscr{F}\{I\} \cdot W\} \ . \tag{2.3}$$

That is, deconvolution can be considered as a filtering operation.

However, from a practical standpoint, treating the imaging equation as being solely due to the convolution of the object with the PSF for the system, is too simplistic: there is always some contamination (e.g. photon shot noise, camera read noise). A better model for an observed image is thus:

$$I = O * PSF + C \tag{2.4}$$

or, equivalently,

$$\mathscr{F}\{I\} = \mathscr{F}\{O\} \cdot OTF + \mathscr{F}\{C\} \ , \tag{2.5}$$

where C represents noise contamination. Now when we use a simple inverse filter we can see that we have a term $\mathscr{F}\{C\}/OTF$ which can cause noise amplification in regions where this ratio swamps $\mathscr{F}\{O\}$.

We can circumvent this issue by defining a modified inverse filter

$$W = \frac{OTF^*}{|OTF|^2 + \Psi} \tag{2.6}$$

which is known as the *Wiener filter*. Here $*$ denotes complex conjugate, and the function Ψ is a measure of the noise-to-signal ratio as a function of spatial frequency. That is, Ψ is an estimate of $|\mathscr{F}\{C\}/\mathscr{F}\{O\}|^2$.

Thus the filter defined by equation (2.6) satisfies

$$\begin{aligned} W &\simeq 1/OTF && \text{when} \ \ |OTF| \gg \Psi \\ &\simeq OFT^*/\Psi && \text{when} \ \ |OTF| \ll \Psi \end{aligned} \tag{2.7}$$

We note that the presence of noise contamination in the imaging equation makes it impossible to obtain an exact solution to the deconvolution problem. [1]

When using a Wiener filter approach with a time series of k images of speckle data (i.e. speckle holography), we have

$$O = \mathscr{F}^{-1}\left(\frac{\sum_k \mathscr{F}\{I_k\} \cdot OTF_k^*}{\sum_k \left(|OTF_k|^2 + \Psi_k\right)}\right) \ . \tag{2.8}$$

Blind deconvolution

In the previous section we assumed that we had perfect knowledge of the PSF. This is rarely the case. In some scenarios we may have a reasonable estimate of the PSF (for example, long-exposure imagery) whilst in others we may have only poor (or no) knowledge. The latter is usually the case when dealing with imagery obtained through atmospheric turbulence. Fortunately, for imagery where we have poor knowledge of the PSF we can turn to the technique of blind deconvolution. Blind deconvolution provides a way to estimate both the original object and the PSF from just the blurred image. There are many algorithms for solving blind image deconvolution but here will focus on iterative blind deconvolution.

---

[1] IMAGE: NOISY BLURRED IMAGE; SIMPLE INVERSE FILTERED IMAGE; WIENER FILTERED IMAGE

In blind deconvolution we are solving for both the object and the PSF given just a single noisy image. This make blind deconvolution an ill-posed inverse problem and in order to make any headway in solving the problem we need to invoke prior information on the physics of the imaging system. There are many forms of prior information that have been invoked in blind deconvolution but probably the most powerful prior information is that for incoherent imaging system we are measuring intensities which are represented by positive functions. That is, both the object and the PSF are positive functions.

The basic iterative blind deconvolution approach we will examine (there are many variations) is given by 4 steps:

1. Our goal the values of the pixels in $\hat{O}$ and $P\hat{S}F$ that minimize the cost function (also called objective function in computer science):

$$\epsilon(O, PSF) = \sum_{x,y} w(x,y) \left(I(x,y) - \hat{I}(x,y)\right)^2 \ , \tag{2.9}$$

   where $\hat{I} = \hat{O} * P\hat{S}F$ is our model for the observed image I, $\hat{O}$ and $P\hat{S}F$ are our estimates for the truth object and PSF, respectively, and w is a weighting function that reflects the noise statistics in I.

2. Model the object and PSF as 2D pixel arrays in which each pixel is a variable. To ensure positivity on $\hat{O}$ and $P\hat{S}F$ we can reparameterize these quantities as the square of the variables we adjust in the minimization (e.g., $O(x,y) = \beta^2(x,y)$ where $\beta(x,y)$ are the variables).

3. Use iterative gradient-based optimization algorithm to find the solution.[2] Compute the gradient required for the optimization using:

$$d\epsilon/dO(x,y) = -2\ R(x,y) \star PSF(x,y) \tag{2.10}$$

   and

$$d\epsilon/dPSF(x,y) = -2\ R(x,y) \star O(x,y) \tag{2.11}$$

   where the residual $R(x,y) = \left(I(x,y) - \hat{I}(x,y)\right)$.

4. Provide the algorithm with best possible initial estimates for $\hat{O}$ and $P\hat{S}F$.

[2] In this class we will be using a constrained semi-Newton method (constrained: because we enforce positivity, semi-Newton: because it estimates the second derivative to improve convergence speed). The user needs to provide the objective function and its gradient with respect to the minimization parameters. The semi-Newton method estimates the best step to take based on the current gradient, the constraints, and the history of the previous steps. Note that there are other optimization methods possible (all iterative), including gradient-less methods such as Markov Chain Monte-Carlo optimizers (simulated annealing, parallel tempering) and splitting/proximal methods (Half Quadratic Split, ADMM).

We can improve on the modeling of the PSF by invoking another prior: the PSF is a band limited function, i.e. it only has spatial frequencies out to the diffraction limit for the telescope. In this case, instead of a pixel-by-pixel model of the intensity of the PSF, we model the PSF via its pupil function and use a pixel-by-pixel (or Zernike polynomial) representation of the wave front phase in the pupil as the variables (we ignore the effects of scintillation). This approach also has the benefit of requiring fewer parameters to model the PSF.

If a time series of images is available, and the object can be considered stationary over the length of the time series, then we can use multi-frame blind deconvolution (MFBD) and equation 2.9 becomes

$$\epsilon = \sum_k \epsilon_k \tag{2.12}$$

where the index k is over frames. The advantage of MFBD is that the object estimate is common to all frames.

[3]

Diversity imaging methods

We can improve the restoration process by incorporating *prior* information. An example of this is given by the diversity imaging techniques.

1. Phase diversity: In this technique we intentionally apply a known phase perturbation to the incoming wave fronts. The phase perturbation may be a simple defocus, which corresponds to a quadratic variation of phase across the exit pupil and is easy to implement (e.g., simultaneously acquire in- and out-of-focus images), or it may be much more general and complex.

2. Wavelength diversity: An alternative approach for creating phase diversity is to acquire simultaneous images at two or more different wavelengths. Here we can use the prior knowledge that the wavefront phases in the pupil for the different wavelengths all correspond to the same optical path difference (OPD). By solving for the OPD in the pupil we are able to determine the PSFs for the different wavelengths.

3. Aperture diversity: In this approach we simultaneously acquire images using aperture functions with different morphology. One example is to partition the full aperture into a series of annuli and to simultaneously form focal plane images for each annular subaperture. Here the level of turbulence is lowest on the smallest annulus and largest on the largest annulus. This range in turbulence strength in the data can be leveraged to boot-strap a restoration from a high-quality, low-resolution estimate of the target object, to

[3] IMAGE: TRUTH OBJECT, TRUTH PSF; BLURRED, NOISY IMAGE; REVOVERED OBJECT; RECOVERED PSF

IMAGE: show results for MFBD with different numbers of frames a high-resolution restoration. Moreover it allows for restoration of diffraction-limited resolution imagery through high levels of turbulence (e.g. $D/r_0 = 70$).

3

Adaptive Optics

3.1 Wavefront references, wavefront sensing

3.2 AO & MCAO

3.3 Image restoration with AO compensated data

4

Optical interferometry

4.1 Fringes

4.2 Visibilities

Radio versus optical

4.3 Fringe tracking

4.4 Aperture synthesis

4.5 Interferometric imaging

List of Figures

1.1 A Golay array with three sub-apertures. 2

1.2 The first 21 Zernike polynomials, ordered vertically by radial degree and horizontally by azimuthal degree. Each represent a possible mode of the pupil phase. 3

1.3 How to build "by hand" the autocorrelation of a single aperture pupil P. For simplicity we assume here the pupil is uniform and real-valued (zero phase). We apply Equation 1.29 by sliding the pupil on top of itself (here from left to right), and computing the integral of the overlapping surface (hashed). Note that since the pupil is point-symmetric, so is the MTF. We obtain the well-known result that the frequency cutoff is $\frac{D}{\lambda}$. 10

1.4 How to build "by hand" the autocorrelation of a pupil with two sub-apertures. The same principles as for Figure 1.4 apply. This pupil is not point-symmetric, so the shape of the MTF slide depends on the sliding direction: this is obvious since this pupil will provide better resolution along the horizontal axis than the vertical one. 10

List of Tables

1.1 The first few Zernike modes, ordered by Noll index j, normalized so that $\int_0^{2\pi}\int_0^1 Z_j^2\,\rho\,d\rho\,d\theta = \pi$. 4
1.2 Some characteristics of convolution and correlation of two real functions 11
1.3 Some characteristics of autocorrelation of a real function 11

List of Exercises

1.1 Exercices on Zernikes 5
1.2 Exercises on PSF 7
1.3 Exercices on image formation 13