CHAPTER 1
OPTICAL EFFECTS OF ATMOSPHERIC TURBULENCE

J C DAINTY
Blackett Laboratory
Imperial College
London SW7 2BZ - UK

The aim of this Chapter is to provide a *gentle* introduction to the optical effects of atmospheric turbulence on astronomical images. A number of more detailed descriptions should be read in conjunction with this Chapter and those by Roddier[13][14], Fried[5], and Roggemann and Welsh[16] are particularly recommended.

In my view, the only way to really appreciate how atmospheric turbulence affects astronomical images is to make observations and measurements yourself: one quickly learns to appreciate the meaning of words like "random" and "average", and acquire the detailed understanding that is necessary to design and operate adaptive optics systems. Relying on computer simulations is second best and may encourage a false sense of security.

1. A Primer on Statistics and Random Processes

The value f of a random variable F is not precisely predictable: one can only talk of the *probability* of it taking a particular value (for discrete variables such as the value of the throw of a dice, or the number of photons detected), or it lying between two values (for a continuous variable such as the air temperature or an optical wavefront). The *probability density function* $p(f)$ is defined as:

$$p(f)df \equiv \texttt{Prob}[f \leq F < f + df] \tag{1}$$

Rather than specify the whole function $p(f)$, it is sometimes convenient to describe the behaviour of the random variable by a few numbers, the *moments* μ_n and indeed this is sometimes a complete description: for example, a Gaussian distribution (this accurately quantifies the phase distortion introduced by atmospheric turbulence) is *completely* specified by the first two

N. Ageorges and C. Dainty (eds.), Laser Guide Star Adaptive Optics for Astronomy, 1-21.

moments. The first moment μ_1 is the *mean*

$$\mu_1 = \langle f \rangle = \int_{-\infty}^{+\infty} p(f) f \, df \tag{2}$$

and the second moment is

$$\mu_2 = \langle f^2 \rangle = \int_{-\infty}^{+\infty} p(f) f^2 \, df \tag{3}$$

The *variance* σ^2 is defined as

$$\sigma^2 = \mu_2 - \mu_1^2 \tag{4}$$

and the square root of the variance is called the *standard deviation.* The mean and standard deviation are the most common two parameters used to describe a random variable, and random quantities are frequently expressed as `mean` $\pm$ `std dev`. In adaptive optics, the variance σ^2 is often preferred, since when two independent random variables are added, their variances also add: this is useful in "error budgets", where the wavefront phase variances due to the (many) various errors add to give the total wavefront phase variance.

A *random process* $F(x)$ is a sequence of random variables in time or space, or both: the variable x can represent time, space, or both, although in this Chapter we take x to represent just space (distance). The phase of a wave that has propagated through turbulence is a space- and time-varying random process. The value of a random process at one point x is a random variable, and has a probability density function $p(f[x])$, mean $m(x)$ and variance $\sigma(x)$. If the process is *stationary*, then none of these quantities depends on x, that is, the statistics are the same everywhere. Atmospheric turbulence is a stationary process to a good approximation.

The function $p(f)$ is strictly called the *first order* probability density function (pdf) since it only describes the statistics at a single point x. If we wish to describe the spatial or temporal structure of a random process (for example, is it changing slowly or rapidly?) then we need to use a second order probability density function: like the first order pdf, this can be parameterised in terms of its moments, and by far the most important moment is the *covariance function* $C(x')$ which can be written as follows for a stationary real process:

$$C(x') \equiv \langle (f(x) - \langle f \rangle)(f(x + x') - \langle f \rangle) \rangle \tag{5}$$

Note that often $\langle f \rangle = 0$ and also that the variance is given by

$$\sigma^2 = C(x' = 0) = \langle f^2 \rangle - \langle f \rangle^2 \tag{6}$$

As we shall see for the case of atmospheric turbulence, there is sometimes a problem in defining the covariance function (sometimes called the autocorrelation function), and a *structure function* is defined instead:

$$\begin{aligned} D(x') &= \langle (f(x) - f(x+x'))^2 \rangle && (7) \\ &= 2\left[C(0) - C(x')\right] && (8) \end{aligned}$$

Finally, the power spectrum $\Phi(\kappa)$ of a stationary process is the Fourier transform of the covariance function:

$$\Phi(\kappa) = \int_{-\infty}^{+\infty} C(x') \exp\left(-2\pi i \kappa x'\right) dx \tag{9}$$

The power spectrum describes the structure of the process in Fourier, or frequency, space.

2. Kolmogorov Turbulence

The phase of a wave that propagates through turbulence becomes distorted because of the random refractive index fluctuations in the atmosphere. These refractive index fluctuations (and hence the imposed phase fluctuations) are Gaussian to a good approximation. Their spatial statistics were first "derived" by Kolmogorov, using a dimensional argument (see [8] and [4]) and here we simply quote the result. Within a range of separations $\Delta r = |\vec{r_1} - \vec{r_2}|$ ($\vec{r}$ is a three-dimensional vector) that are greater than the *inner scale* ℓ_0 and less than the *outer scale* L_0, the structure function obeys a power law:

$$D_n(\Delta r) = C_N^2 \Delta r^{2/3} \qquad \ell_0 < \Delta r < L_0 \tag{10}$$

This is the famous "two-thirds power law" and we shall find that all the important quantities in atmospheric turbulence have characteristic power laws different amounts of "thirds". The quantity C_N^2 is called the refractive index structure constant: its value depends very strongly with altitude z and so usually we write

$$D_n(\Delta r) = C_N^2(z) \Delta r^{2/3} \tag{11}$$

The units of $C_N^2(z)$ are $\mathrm{m}^{-2/3}$.

The power spectrum of the refractive index fluctuations is given by

$$\Phi_n(\vec{\kappa}) = 0.033 C_N^2(z) |\vec{\kappa}|^{-11/3} \tag{12}$$

within the range $1/L_0 < |\vec{\kappa}| < 1/\ell_0$. When we need to include the effects of finite inner and outer scales, the modifed von Karman spectrum is often used:

$$\Phi(\vec{\kappa}) = \frac{0.033 C_N^2}{(|\vec{\kappa}|^2 + \kappa_0^2)^{11/6}} \exp\left(-\frac{|\vec{\kappa}|^2}{\kappa_m^2}\right) \tag{13}$$

where $\kappa_0 = 2\pi/L_0$ and $\kappa_m = 5.92/\ell_0$.

Some comments on Eqs. 11 and 12 are in order. Eq.11 states that the average refractive index difference between two points increases without bound (with a two-thirds power law) as their separation increases, a result that clearly is unphysical for large separations. The existence of an outer scale makes this result physical, the outer scale being a measure of the value at which this average difference ceases to increase: there are few reliable measurements of the outer scale but it is believed to be on the order of tens of metres[11]. Similarly, examination of Eq.11 reveals that the spectrum approaches infinity as $|\vec{\kappa}| \to 0$: the finite outer scale again prevents this from happening in practice. Eq.13 is widely used in modelling to incorporate values for the inner and outer scales.

When one is interested primarily in the *phase* of a wavefront that has propagated through turbulence (the usual situation in adaptive optics), then the outer scale L_0 is usually of much greater significance that the inner scale ℓ_0, particularly regarding issues such as wavefront tilt. On the other hand, if considering the *intensity* (scintillation) of a wave that has propagated through turbulence, it is the inner scale that is the more significant of the two parameters.

3. Field Correlation and Phase Structure Function

A key concept in adaptive optics is the *wavefront.* A wavefront is found by tracing out an equal optical path (distance $\times$ refractive index) from a source to the region of interest, for example, the entrance pupil of an optical system. For a point source and free space, wavefronts are spherical, and for starlight the distance is so large that for all practical purposes the wavefronts entering the Earth's atmosphere are plane. After propagating through the random refractive index of the atmosphere, the wavefront entering the telescope pupil is random, and it statistics determine the image quality, and also govern how an adaptive optical system might be used to compensate for the distortion.

In this Chapter, we denote the wavefront by $W(\vec{x})$, where $\vec{x}$ is a two-dimensional vector in the plane of the pupil. Since W is defined by an optical path, its units are length (metres), although often ones states the length in units of the wavelength being used. To first order, the wavefront aberration introduced by turbulence is achromatic, i.e. independent of wavelength of light, so that, for example, the number of microns of optical path retardance is the same for both visible and infrared light: however, the number of *wavelengths* of optical path is different — there are fewer waves of retardance at (the longer) infrared wavelengths — and therefore, from this elementary picture, it is clear that the distortion experienced by infrared light is less

serious than that experienced by visible light. It immediately follows that optical degradation at IR wavelengths should be easier to compensate for. It should be noted that the refractive index of air is not *quite* the same for all wavelengths, but that like any of material it varies with wavelength, that is, there is dispersion. As described in a later Chapter, this dispersion can be used to advantage in a polychromatic laser guide star to determine the absolute tilt due to atmospheric turbulence.

Although the basic quantity of interest is the wavefront aberration, it is more usual to deal with the *phase fluctuation* $\phi(\vec{x})$ where

$$\phi(\vec{x}) = \frac{2\pi}{\lambda} W(\vec{x}) \tag{14}$$

where λ is the wavelength. Note that this phase distortion is specified between $-\infty$ and $+\infty$, whereas optically, at a given wavelength, we cannot distinguish between ϕ and $\phi \pm 2n\pi$, where n is any integer. We shall denote the 2π wrapped phase as $\phi|_{\mathrm{Mod}2\pi}$. Clearly the phase distortion ϕ or $\phi|_{\mathrm{Mod}2\pi}$ is also a strong function of wavelength.

The complex amplitude $U(\vec{x})$ of a plane wave of wavelength λ that has propagated through turbulence is given by

$$U(\vec{x}) = |U(\vec{x})| \exp i\phi(\vec{x}) = |U(\vec{x})| \exp\left(\frac{2\pi i}{\lambda} W(\vec{x})\right) \tag{15}$$

The problem of finding the properties of the complex amplitude $U(\vec{x})$ is quite involved: we basically have to solve the wave equation in a three-dimensional random medium. The first order statistics are simple: $|U(\vec{x})|$ is a log-normal process and ϕ (or W) is Gaussian. If the turbulence is not too strong — usually the case in astronomy — then a detailed analysis (the clearest derivation is given by Roddier [13]) shows that the covariance function of the complex amplitude for a wave that has propagated through Kolmogorov turbulence is given by

$$\begin{aligned} C(\vec{x}') &= \langle U(\vec{x}) U^*(\vec{x} + \vec{x}') \rangle & (16) \\ &= \exp\left(-\frac{1}{2} D_\phi(\vec{x}')\right) & (17) \end{aligned}$$

where $\langle U \rangle = 0$ and $D_\phi(\vec{x}')$ is the *phase structure function*. This in turn is given by

$$\begin{aligned} D_\phi(\vec{x}') &= 2.91\ k^2 (\cos\gamma)^{-1} |\vec{x}'|^{5/3} \int_0^\infty C_N^2(z) dz & (18) \\ &= 6.88 \left(\frac{|\vec{x}'|}{r_0}\right)^{5/3} & (19) \end{aligned}$$

where $k = 2\pi/\lambda$ and r_0, the Fried parameter, is given by

$$r_0 = \left(\frac{2.91}{6.88}k^2 \cdot (\cos\gamma)^{-1} \int_0^\infty C_N^2(z)dz\right)^{-3/5} \tag{20}$$

In these equations, γ is the zenith angle (angle of observation measured from the zenith).

It is clear from Eqs.17 and 19 that the *Fried parameter* r_0 is a crucial quantity, since if this is known then both the phase structure function and field correlation are also determined. Fried defined this quantity in his 1965 paper [2], and r_0 remains the single most important parameter describing the quality of a wave that has propagated through turbulence. The Fried parameter also has two physical interpretations. First, it is the aperture over which there is approximately one radian of root-mean-square phase aberration. Second, it is the aperture which has the "same resolution" (as defined by Fried) as a diffraction-limited aperture in the absence of the turbulence.

A typical value for r_0 in the visible region of the spectrum lies in the range of 10–20cm: this means therefore that there is approximately one radian of *rms* phase error over a (say) 10cm aperture and that the image resolution, in the presence of this amount of turbulence, is equivalent to that of a *diffraction-limited* optical system (no turbulence) of 10cm aperture. In other words, if the "seeing" is 10cm at the Keck 10m telescope, the image resolution is no better than that given by a 10cm amateur telescope!

If one measures r_0 (using, for example, a differential image motion monitor [17]), it is often found to be a highly fluctuating quantity from minute to minute and can easily vary by a factor of two over a few minutes. It is a feature of human nature that astronomers always remember the best moments of seeing!

There are two important functional dependences of r_0 that should be noted: first, r_0 is a simple integral over the $C_N^2(z)$ profile:

$$r_0 \propto \left(\int_0^\infty C_N^2(z)dz\right)^{-3/5} \tag{21}$$

We shall see that all the important atmospheric parameters for adaptive optics are integrals (some are weighted integrals) over $C_N^2(z)$, as thus the profile of the refractive index structure constant is a very important quantity. Second, r_0 is proportional to the six-fifths power of the wavelength:

$$r_0 \propto \lambda^{6/5} \tag{22}$$

Table 1 gives some typical values of the Fried parameter at several wavelengths. The wavelength dependence is of crucial importance, since the ratio of $\mathcal{D}/r_0$ is important.

λ =0.55μm	λ =1.2μm	λ =1.6μm	λ =2.2μm
10	25	36	53
15	38	54	79
20	50	72	106

TABLE 1. Values of the Fried parameter r_0, in cm, for the given wavelengths, assuming a six-fifths power law dependence.

The power spectrum of the phase fluctuation corresponding to the structure function of Eq.19 is

$$\Phi_\phi(\vec{\kappa}) = \frac{0.023}{r_0^{5/3}} |\vec{\kappa}|^{-11/3} \tag{23}$$

where $\vec{\kappa}$ is a two-dimensional frequency vector.

4. Strehl Ratio, Marechal Criterion

In this section, we address the relationship between the phase aberration and image quality. This is traditionally quantified, in a high quality imaging system, by the *Strehl ratio* $\mathcal{S}$, defined as the ratio of the central intensities of the aberrated point spread function and the diffraction-limited point spread function:

$$\mathcal{S} = \frac{I(0,0)_{\text{aberrated}}}{I(0,0)_{\text{d-l}}} \tag{24}$$

where $I(\xi, \eta)$ is the intensity point spread function, and (ξ, η) are image plane coordinates (both equal to zero in this case).

For *small* arbitrary aberrations, the Strehl ratio is related to the *variance* of the phase aberration by

$$\mathcal{S} \approx 1 - \sigma_\phi^2 \qquad \sigma_\phi^2 \ll 1 \tag{25}$$

A system is said to be "well-corrected", i.e. close to diffraction-limited in practical terms, when $\mathcal{S}$ $\geq$0.8, the equality being called the Strehl or Marechal criterion. At the Marechal limit,

$$\sigma_\phi^2 = \left(\frac{2\pi}{\lambda}\right)^2 \sigma_W^2 = 0.2 \qquad \text{rad}^2 \tag{26}$$

corresponding to an rms wavefront aberration of $\sigma_W \approx \lambda/14$ or a wavefront variance $\sigma_W^2 \approx \lambda^2/200$. Note that the Marechal residual phase value of 0.2

rad^2 is smaller than the residual phase of approximately one rad^2 in the definition of the Fried parameter r_0. As we shall see in the next section, the diameter of telescope required to achieve a Strehl ratio of 0.8 (in the absence of any adaptive optics), is approximately $0.4r_0$ (i.e. significantly smaller than r_0).

If the phase aberration is *not* small, then the Strehl ratio depends on the specific form of the aberration. However, for a random Gaussian aberration, such as that given by atmospheric turbulence, provided that the telescope diameter $\mathcal{D}$ is very much greater than r_0, it can be shown that

$$\mathcal{S} \approx \exp\left(-\sigma_\phi^2\right). \tag{27}$$

This result is a property of Gaussian random processes. Since it subsumes the previous expression, we will take Eq.27 as the definitive form of the Strehl ratio in terms of the variance of the phase aberration. Naïvely, one might expect the goal of an adaptive optics system would to reach the Marechal limit of $\mathcal{S}$ =0.8 by reducing the residual phase variance, after correction, to 0.2 rad^2: in practical terms, this is a rather ambitious goal, at least for imaging through the atmosphere.

5. Zernike Expansion of Kolmogorov Turbulence

The Zernike polynomials $Z_j(\rho,\theta)$ are an orthogonal expansion over the unit circle, and have a long tradition of use in classical optical aberration analysis. They have been adopted by the AO community as the *de facto* standard rightly or wrongly (probably the latter). They are defined as [10],[9]:

$$\begin{aligned}
Z_j &= \sqrt{(n+1)}\;\; R_n^0(\rho), && m=0 \\
Z_{\mathrm{even}\ j} &= \sqrt{(n+1)}\;\; R_n^m(\rho)\sqrt{2}\cos m\theta && m\neq 0 \\
Z_{\mathrm{odd}\ j} &= \sqrt{(n+1)}\;\; R_n^m(\rho)\sqrt{2}\sin m\theta && m\neq 0 \qquad (28)
\end{aligned}$$

where

$$R_n^m(\rho) = \sum_{s=0}^{(n-m)/2} \frac{(-1)^s (n-s)!}{s![(n+m/2-s]![n-m/2-s]!}\rho^{(n-2s)} \tag{29}$$

The orthogonality of the Zernike polynomials is expressed by

$$\int_0^R \int_0^{2\pi} \mathcal{W}(\rho) Z_j(\rho,\theta) Z_k(\rho,\theta)\rho d\rho d\theta = \delta_{jk} \tag{30}$$

where the aperture weighting function $\mathcal{W}(\rho)$ is defined by

$$\begin{aligned}
\mathcal{W}(\rho) &= 1/\pi && r\leq 1 \\
&= 0 && r>1.
\end{aligned}$$

j	n	m	**Zernike Polynomial**	**Name**
1	0	0	1	Constant
2	1	1	$2\rho\cos\theta$	Tilt
3	1	1	$2\rho\sin\theta$	Tilt
4	2	0	$\sqrt{3}(2\rho^2-1)$	Defocus
5	2	2	$\sqrt{6}\rho^2\sin 2\theta$	Primary Astigmatism
6	2	2	$\sqrt{6}\rho^2\cos 2\theta$	Primary Astigmatism
7	3	1	$\sqrt{8}(3\rho^3-2\rho)\sin\theta$	Primary Coma
8	3	1	$\sqrt{8}(3\rho^3-2\rho)\cos\theta$	Primary Coma
9	3	3	$\sqrt{8}\rho^3\sin 3\theta$	
10	3	3	$\sqrt{8}\rho^3\cos 3\theta$	
11	4	0	$\sqrt{5}(6\rho^4-6\rho^2+1)$	Primary Spherical
12	4	2	$\sqrt{10}(4\rho^4-3\rho^2)\cos 2\theta$	
13	4	2	$\sqrt{10}(4\rho^4-3\rho^2)\cos 2\theta$	
14	4	4	$\sqrt{10}\rho^4\cos 4\theta$	
15	4	4	$\sqrt{10}\rho^4\sin 4\theta$	
16	5	1	$\sqrt{12}(10\rho^5-12\rho^3)\cos\theta$	
17	5	1	$\sqrt{12}(10\rho^5-12\rho^3)\sin\theta$	
18	5	3	$\sqrt{12}(5\rho^5-4\rho^3)\cos 3\theta$	
19	5	3	$\sqrt{12}(5\rho^5-4\rho^3)\sin 3\theta$	
20	5	5	$\sqrt{12}\rho^5\cos 5\theta$	
21	5	5	$\sqrt{12}\rho^5\sin 5\theta$	
22	6	0	$\sqrt{7}(20\rho^6-30\rho^4+12\rho^2-1$	Secondary Spherical

TABLE 2. First 22 Zernike polynomials.

Table 2 lists the first 22 Zernike polynomials, with the numbering convention that the first one is j =1.

Figure 1 shows the form of the first few Zernike polynomials.

Any wavefront or wavefront phase can be expanded in terms of Zernike polynomials. For a circular pupil of radius R, the expansion is

$$\phi(R\rho,\theta)=\sum_j a_j Z_j(\rho,\theta) \tag{31}$$

where the coefficients a_j are given by

$$a_j=\int_0^R\int_0^{2\pi}\mathcal{W}(\rho)\phi(R\rho,\theta)Z_j(\rho,\theta) \tag{32}$$

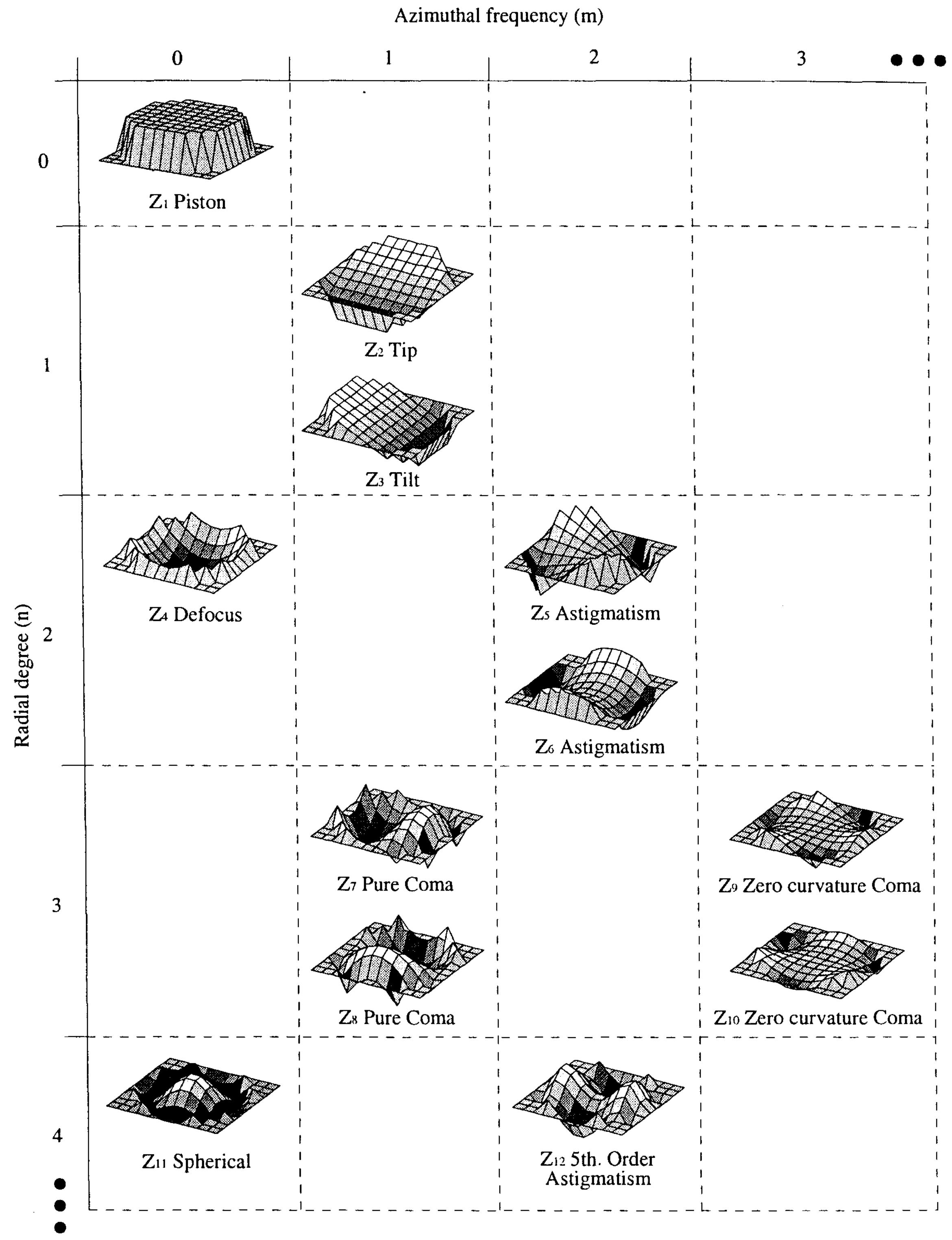


Figure 1. First 12 Zernike polynomials (courtesy of P Negrete-Regagnon).

This last equation states that the a_j are linear combinations of the pupil phase $\phi(\rho, \theta)$ and therefore the a_j are Gaussian random variables if the

phase is Gaussian.

Noll [12] was the first person to apply Zernike polynomials to the description of atmospheric turbulence. He showed that the Zernike coefficients are weakly correlated (Eq.25 of [12]), the correlation existing only when both m and the parity are different. He also calculated the residual mean square phase error that resulted when the first J Zernike terms are corrected and the results of this calculation are given below.

The wavefront phase resulting from the summation of the first J Zernike terms is

$$\phi_C(R\rho,\theta) = \sum_{j=1}^{J} a_j Z_j(\rho,\theta) \tag{33}$$

so that the mean squared phase error σ_ϕ^2, or to use Noll's terminolgy Δ_J, when these J terms are exactly corrected is

$$\begin{aligned}\sigma_\phi^2 &= \Delta_J = \int_{-\infty}^{+\infty} W(\rho,\theta)\langle[\phi(R\rho,\theta) - \phi_C(R\rho,\theta)]^2\rangle \rho d\rho d\theta \\ &= \langle\phi^2\rangle - \sum_{j=1}^{J}\langle|a_j|^2\rangle\end{aligned}$$

As noted earlier, the infinity of $\langle\phi^2\rangle$ is contained in the j =1 term (piston), so that Δ_1 (correction for piston) is finite. Table 3 lists the residual phase variances for J =1 to J =21, for a telescope of diameter $\mathcal{D}$ and a Fried parameter r_0.

It can now be seen from Table 3 that if $\mathcal{D} = r_0$, then then piston-removed wavefront variance is 1.03 rad^2: to achieve a Strehl ratio of 0.8, the residual phase variance would have to equal 0.2 rad^2, implying $\mathcal{D} \approx 0.4 \times r_0$. Note the large effect on the residual variance of the correction of piston + tip + tilt (i.e. Δ_3 in Table 3): approximately 87% of the wavefront variance is contained in these terms, and it is tempting to think that correction of only these terms might give significant increase in the Strehl value. Obviously the degree of correction required to achieve a certain Strehl value depends very strongly upon $\mathcal{D}/r_0$: using the approximate expression given by Noll,

$$\Delta_J \approx 0.29 J^{-\frac{\sqrt{3}}{2}} \left(\frac{\mathcal{D}}{r_0}\right)^{\frac{5}{3}} \qquad \text{rad}^2 \qquad J \gg 1 \tag{34}$$

we obtain the relationship

$$\begin{aligned}\mathcal{S} &= \exp(-\Delta_J) && (35)\\ &\approx 1 - \Delta_J \qquad \Delta_J \ll 1 && (36)\end{aligned}$$

so that we can plot graphs of the required number of Zernike polynomials that have to be corrected as function of $\mathcal{D}/r_0$ to achieve given Strehl values.

j	Squared error	j	Squared Error
1	1.030	12	0.0352
2	0.582	13	0.0328
3	0.134	14	0.0304
4	0.111	15	0.0279
5	0.0880	16	0.0267
6	0.0648	17	0.0255
7	0.0587	18	0.0243
8	0.0525	19	0.0232
9	0.0463	20	0.0220
10	0.0401	21	0.0208
11	0.0377		

TABLE 3. The residual phase variance (in rad^2) when the first J Zernike coefficients of Kolmogorov turbulence are corrected exactly (after Noll). The values given are for $\mathcal{D} = r_0$ and scale as $(\mathcal{D}/r_0)^{5/3}$.

This is done in Fig. 2, for Strehl ratios of 0.1, 0.4 and 0.8. For example, if $\mathcal{D}/r_0$ =40 (a typical value for a 4m diameter telescope in the visible part of the spectrum in average seeing), then >1000 Zernike terms must be corrected, if the Marechal criterion is to be reached! However, if the wavelength is increased to 2.2μm, the number of Zernikes to be corrected reduces dramatically to about 70.

Of course, an adaptive optical system does not work by compensating for the lowest J Zernike terms exactly: in practice, it is not terms in a Zernike expansion that are corrected, and anyway *nothing* is corrected exactly. However, the above calculations are useful as they give an idea of the "degrees-of-freedom" required to obtain good correction and give an idea of the overall magnitude of the correction problem.

As remarked above, the Zernike coeeficients are weakly correlated, and therefore the Zernike expansion is *not* the best expansion to use for Kolmogorov turbulence: the optimum expansion is the Karhunen-Loevé expansion. Roddier [15] shows that these can be expressed in terms of Zernike polynomials and a comparison of the two can be found in the book by Roggemann and Welsh [16].

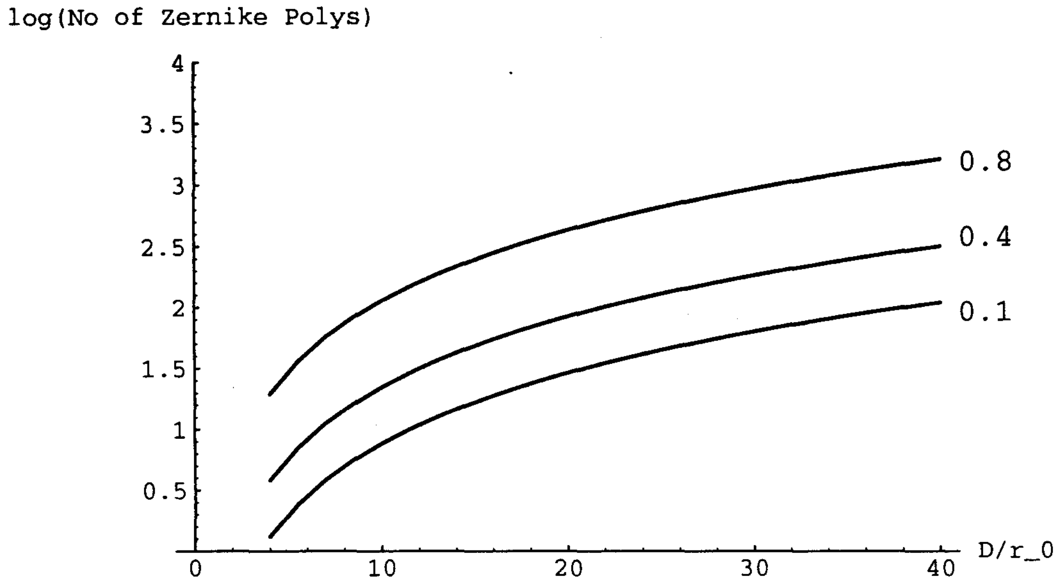


Figure 2. The logarithm (base 10) of the number of Zernike polynomials that must be precisely compensated for, in order to reach Strehl ratios of 0.8 (upper), 0.4 (middle) and 0.1 (lower), as a function of $\mathcal{D}/r_0$.

6. Angle-of-arrival Statistics

Determining the (average) slope of a wavefront is important for two reasons in adaptive optics. First, the correction of average slope ("tip" and "tilt") is, in itself, an important correction, and one can view it as the first step in the implementation of an AO system, although as we saw in the last section, it is rarely sufficient. Second, wavefront sensing is often done by sensing the tilt over sub-apertures, for example, the lenslets in a Shack-Hartmann sensor. Noting the relationship between the wavefront and the phase:

$$W(x,y) = \frac{\lambda}{2\pi}\phi(x,y)$$

we define the slopes along the x and y directions as:

$$\alpha(x,y) \equiv \frac{\partial}{\partial x}W(x,y) = \frac{\lambda}{2\pi}\frac{\partial}{\partial x}\phi(x,y) \tag{37}$$

$$\beta(x,y) \equiv \frac{\partial}{\partial y}W(x,y) = \frac{\lambda}{2\pi}\frac{\partial}{\partial y}\phi(x,y) \tag{38}$$

If $\phi(x,y)$ is Gaussian, then $\alpha(x,y)$ and $\beta(x,y)$ are also Gaussian. This means that the image moves around the image plane with a Gaussian probability distribution.

Using the differentiation property of Fourier transforms, the power spectra of the slopes can be written as:

$$\Phi_\alpha(\vec{\kappa}) = \lambda^2 \kappa_x^2 \Phi_\phi(\vec{\kappa}) \tag{39}$$
$$\Phi_\beta(\vec{\kappa}) = \lambda^2 \kappa_y^2 \Phi_\phi(\vec{\kappa}) \tag{40}$$

where κ_x and κ_y are the components of κ, $\Phi_\phi(\vec{\kappa})$ is the power spectrum of the phase (Eq.23) and $\Phi_{\alpha,\beta}(\vec{\kappa})$ are power spectra of the wavefront slopes. The variance of the slope along say x is therefore

$$\sigma_\alpha = \lambda^2 \int_{\text{bandpass}} \kappa_x^2 \Phi_\phi(\vec{\kappa}) d\vec{\kappa}$$

and the total variance is

$$\sigma_{\alpha\beta}^2 = \lambda^2 \int_{\text{bandpass}} |\kappa|^2 \Phi_\phi(\vec{\kappa}) d\vec{\kappa} \tag{41}$$

A rigorous analysis of this problem was carried out by Fried [2] and the result is:

$$\sigma_{\alpha\beta}^2 = 0.36 \lambda^2 \mathcal{D}^{-1/3} r_0^{-5/3} \tag{42}$$
$$= 0.36 \left(\frac{\lambda}{\mathcal{D}}\right)^2 \left(\frac{\mathcal{D}}{r_0}\right)^{5/3} \tag{43}$$

Three comments on Eq.43 are worthwhile. The apparent wavelength dependence does not exist, as it should be remembered that r_0 is proportional to $\lambda^{6/5}$. Secondly, note the weak dependence on the telescope diameter, which is also unphysical as $\mathcal{D} \to 0$: the finite *inner scale* takes care of this problem. Finally, note that $(\lambda/\mathcal{D})$ is approximately the angular width of the diffraction-limited image, so that the standard deviation of the image motion is approximately $0.6(\mathcal{D}/r_0)^{5/6}$ in units of the width of the *diffraction-limited* point spread function.

Consider an example of a $\mathcal{D}$ =1m telescope, wavelength 0.5μm and r_0 =10cm (typical seeing in the mid-visible): in that case,

$$\sigma_{\alpha\beta}^2 \approx 4.2 \times 10^{-12} \quad \text{rad}^2$$
$$\approx 0.18 \quad (\text{arcsec})^2$$

or a standard deviation of approximately 0.42 arcsec.

Repeating this calculation for $\mathcal{D}$ =10cm, a value appropriate for the lenslets of a Shack-Hartmann array in the visible, gives a standard deviation of approximately 0.62 arcsec.

7. Long- and Short-exposure Transfer Functions

The performance of an optical imaging system is often quantified by either a *point spread function* or a *transfer function.* In the present case of astronomical imaging through turbulence, we are interested in the imaging of spatially incoherent objects, so it is the *intensity* point spread function that is the basic quantity of interest, or alternatively, its Fourier transform, the Optical Transfer Function (OTF).

Over a time that is short compared to the coherence time of the atmosphere (for example, less than a millisecond in the visible region), the object intensity $O(\xi,\eta)$ and image intensity $I(\xi,\eta)$ are related by

$$\begin{aligned} I(\xi,\eta) &= \int\int_{-\infty}^{+\infty} O(\xi',\eta')\mathcal{P}(\xi'-\xi,\eta'-\eta)\; d\xi' d\eta' \\ &= O(\xi,\eta)\odot\mathcal{P}(\xi,\eta) \end{aligned} \tag{44}$$

where $\mathcal{P}(\xi,\eta)$ is is intensity point spread function and $\odot$ represents convolution. In Fourier space this can be represented as

$$\tilde{I}(u,v)=\tilde{O}(u,v)\cdot\mathcal{H}(u,v) \tag{45}$$

where the tilde denotes a Fourier transform and $\mathcal{H}(u,v)$ is the Fourier transform of $\mathcal{P}(\xi,\eta)$; $\mathcal{H}(u,v)$ is called the instantaneous optical transfer function (OTF).

The instantaneous OTF $\mathcal{H}(u,v)$ is given by the normalised autocorrelation of the pupil function:

$$\mathcal{H}(u,v)=P(x,y)\star P(x,y) \tag{46}$$

where $\star$ denotes spatial autocorrelation, distances in the pupil x,y are related to spatial frequencies u,v of the image by

$$x=\lambda f_T u \qquad\qquad y=\lambda f_T v$$

f_T is the focal length of the telescope and the pupil function is defined by

$$P(x,y)=\exp\left(ikW(x,y)\right)$$

Both the instantaneous point spread function $\mathcal{P}(\xi,\eta)$ and the instantaneous OTF $\mathcal{H}(u,v)$ are random functions because of the random pupil function. The simple *average* point spread function is obtained by summing all the point spread functions in the normal way, as would happen in a long-exposure image. We therefore denote it $\langle\mathcal{P}(\xi,\eta)\rangle_{LE}$ and the corresponding long-exposure OTF by $\langle\mathcal{H}(u,v)\rangle_{LE}$. However, if each instantaneous point spread function were *centroided* prior to averaging, (e.g. by an adaptive

tip-tilt mirror), then one would obtain an average point spread function called the short-exposure point spread function, $\langle \mathcal{P}(\xi, \eta) \rangle_{SE}$ and the corresponding short-exposure OTF is denoted $\langle \mathcal{H}(u, v) \rangle_{SE}$.

The long-exposure transfer function is given by [7]:

$$\langle \mathcal{H}(u, v) \rangle_{LE} = C(\lambda f_T u, \lambda f_T v) T_o(u, v) \tag{47}$$

where $T_o(u, v)$ is the optical transfer function of the telecope optical system, and the quantity $C(\lambda f_T u, \lambda f_T v)$ was defined in Eq.(16): it is the correlation of the complex amplitude across the telescope aperture. Substituting from Eq.(19),

$$\langle \mathcal{H}(u, v) \rangle_{LE} = T_o(\vec{u}) \cdot \exp\left(-3.44 \left(\frac{\lambda f_T |\vec{u}|}{r_0} \right)^{5/3} \right) \tag{48}$$

The two parts of the long-exposure OTF are plotted in Fig. 3 for the case where the optical system (telescope) is diffraction-limited — this is a rather unlikely scenario — and $\mathcal{D}/r_0 \approx 10$. Notice the enormous loss in resolution, a factor of 10 (i.e. $\mathcal{D}/r_0$) approximately.

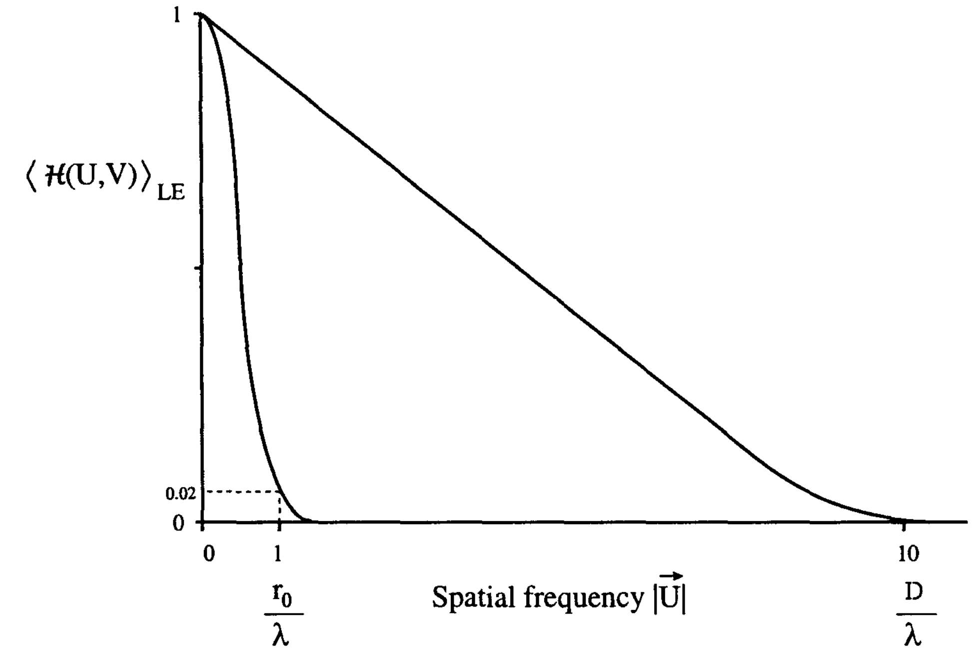


Figure 3. The two terms that make up the long-exposure transfer function $\langle \mathcal{H}(u, v) \rangle_{LE}$ for a value of $\mathcal{D}/r_0 \approx 10$.

The short-exposure transfer function was first derived by Fried [3] (again!), using the results of §2.6, specifically Eq.(43):

$$\langle \mathcal{H}(u, v) \rangle_{SE} = T_o(\vec{u}) \cdot \exp\left(-3.44 \left(\frac{\lambda f_T |\vec{u}|}{r_0} \right)^{5/3} \right) \left[1 - \left(\frac{\lambda f_T |\vec{u}|}{\mathcal{D}} \right)^{1/3} \right] \tag{49}$$

The extra term in Eq.(49) means that, for given seeing conditions, there is an optimum telescope diameter $\mathcal{D}$ for maximum resolution. This can be seen as follows. First we define the resolution $\mathcal{R}$ (a *bandwidth*) by

$$\mathcal{R} \equiv \int_{-\infty}^{+\infty} \langle \mathcal{H}(\vec{u}) \rangle \, d\vec{u} \tag{50}$$

Now consider a fixed value of r_0: then a plot of $\mathcal{R}$ versus telescope diameter $\mathcal{D}$ looks as shown on the lower curve in Fig. 4, i.e. it increases to an asymptotic value denoted by $\mathcal{R}_\infty$. If we now plot the ratio of $\mathcal{R}/\mathcal{R}_\infty$ for the short-exposure case, as shown in the upper curve of Fig. 4, then we see that a maximum of $\mathcal{R}/\mathcal{R}_\infty \approx 3.5$ occurs at $\mathcal{D}/r_0 \approx 3.8$.

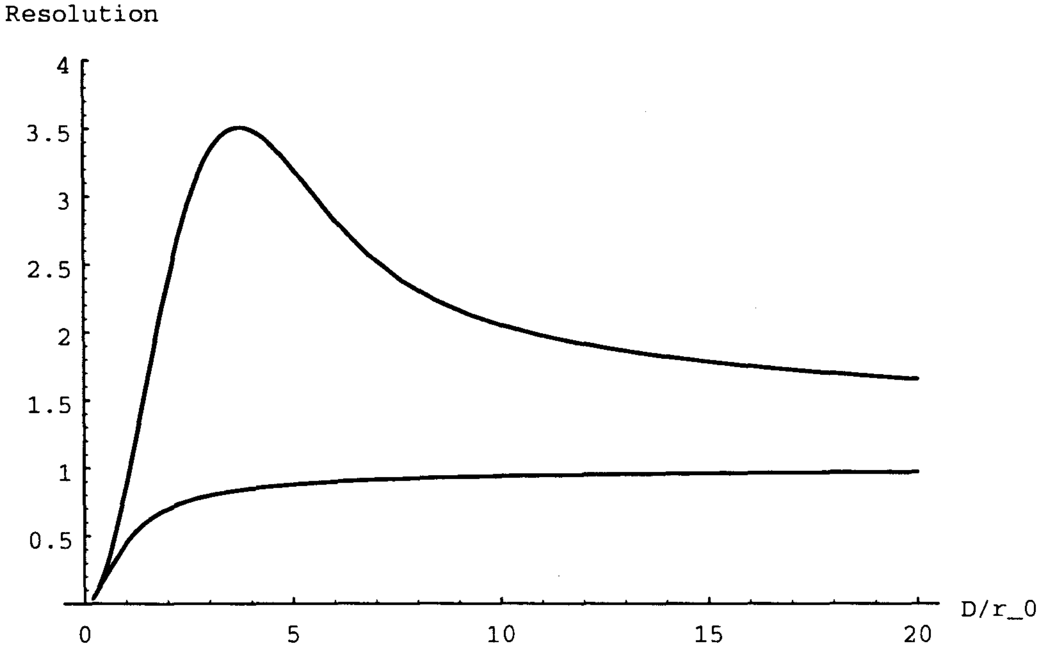


Figure 4. The resolution (bandwidth) $\mathcal{R}$ normalised to the value $\mathcal{R}_\infty$ at $\mathcal{D}/r_0 \to \infty$, for long-exposure (lower curve) and short-exposure (upper curve) imaging. Note the peak value of $\mathcal{R}/\mathcal{R}_\infty \approx 3.5$ at $\mathcal{D}/r_0 \approx 3.8$.

Thus, for a simple tip-tilt system, the optimum value of $\mathcal{D}/r_0$ is approximately 3.8, and an improvement in resolution of between three and four is obtained. However, beware: the improvement in resolution depends strongly on it how resolution is defined, and the value of ≈ 3.5 above is only for the particular definition of $\mathcal{R}$.

8. Temporal Behaviour of Atmospheric Turbulence

A description of the temporal behaviour of the phase in the telescope pupil is rather complicated. The basic model is that there are several strong

layers of turbulence in the atmosphere, and each layer (height z) is moving at some velocity perpendicular to the Earth's surface $\vec{v}_\perp(z)$. If the layer simply moves rigidly, without changing its refractive index distribution, then that layer contributes a phase in the pupil which also just moves rigidly across the pupil: this is called the Taylor hypothesis. In reality, of course, the layer evolves in a statistical fashion as it moves. The result of several layers moving at different velocities gives rise to a complicated phase evolution in the telescope pupil.

A good discussion of temporal effects is given by Conan et al [1] and here we simply cite some useful results. Assuming the Taylor hypothesis to be true, and assuming there to be just *one* layer moving at velocity $\vec{v}_\perp$, then the temporal power spectrum $\Phi_\phi(f)$ of the phase in the telescope pupil is given by:

$$\Phi_\phi(f) \propto (C_N^2\ dz)\frac{1}{|\vec{v}_\perp|}\left(\frac{f}{|\vec{v}_\perp|}\right)^{-8/3} \tag{51}$$

That is, there is -8/3 power law dependence of the phase on the frequency f at any point in the pupil.

In wavefront sensors, we are often interested in the time-dependence of the spatial derivative of the phase, such as $\partial\phi/\partial x$. For a single point (no spatial integration effects) the power spectrum of the phase derivative $\Phi_{\partial\phi/\partial x}(f)$ is:

$$\Phi_{\partial\phi/\partial x}(f) \propto (C_N^2\ dz)\frac{1}{|\vec{v}_\perp|}\left(\frac{f}{|\vec{v}_\perp|}\right)^{-2/3} \tag{52}$$

That is, there is -2/3 power law dependence of the phase derivative on the frequency f at any *point* in the pupil. In reality, we find the average slope over a dimension d, and the results become very much more involved and vary depending upon the sub-aperture shape (circular or square) and the wind direction with respect to the direction of the phase derivative. At very low temporal frequencies, there is the -2/3 power law dependence of Eq.(52), since these frequencies arise from low spatial frequencies which are barely affected by spatial integration. At the high frequency limit, Conan et al [1] found a power law between -11/3 and -14/3 for a square aperture, depending whether the derivative is parallel or perpendicular to the wind velocity.

For adaptive optics, a useful parameter is the Greenwood frequency [6], f_G, defined for an arbitrary $C_N^2(z)$ and $\vec{v}_\perp(z)$ profile as

$$f_G = \left[0.102\left(\frac{2\pi}{\lambda}\right)^2 (\cos\gamma)^{-1}\int_0^\infty C_n^2(z)\,(\vec{v}_\perp(z))^{5/3}\,dz\right]^{\frac{3}{5}} \tag{53}$$

For a single layer of turbulence moving at velocity $\vec{v}_\perp$, using Eq.(20), this reduces to simply

$$f_G = 0.43 \frac{|\vec{v}_\perp|}{r_0} \tag{54}$$

As a simple example, for r_0=10cm (typical value in visible) and $|\vec{v}_\perp|$=10ms^{-1}, then f_G ≈40Hz, this value falling to only ≈8Hz at a wavelength of 2.2μm in the infrared. Typical values from real multilayer turbulence are likely to be higher than these values.

The practical utility of the Greenwood frequency is that, in a first order loop compensator, the residual phase error is related to the Greenwood frequency by

$$\sigma^2_{\text{closed-loop}} = \left(\frac{f_G}{f_{3dB}}\right)^{5/3} \qquad \text{rad}^2 \tag{55}$$

where f_{3dB} is the 3dB closed-loop bandwidth of the compensator.

9. Angular Anisoplanatism

Our final topic in atmospheric turbulence is the subject of angular isoplanatism. Figure 5 shows the problem:

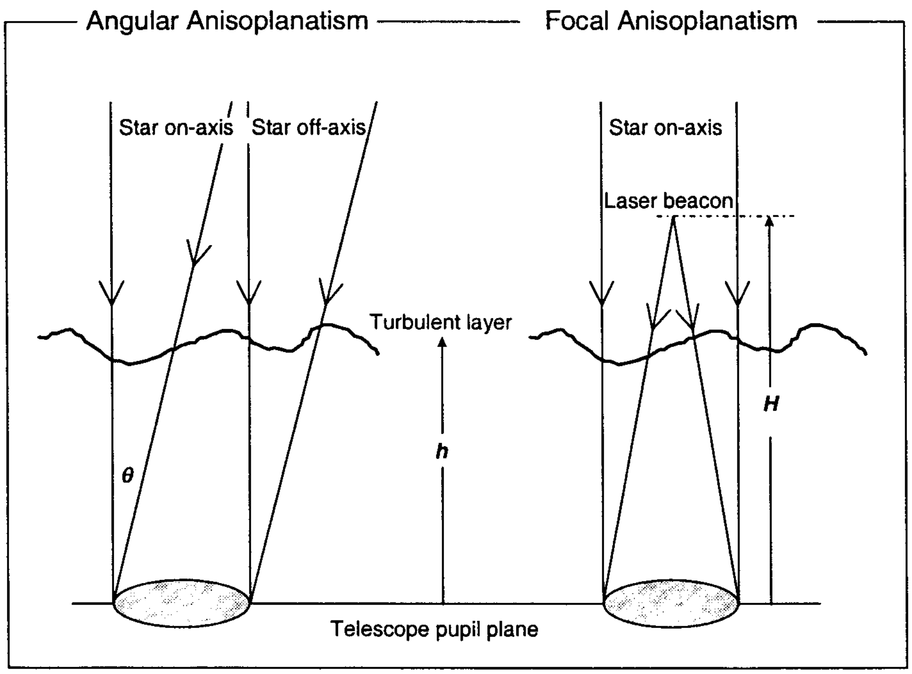


Figure 5. Angular isoplanatism (left) and focus anisoplanatism, or the cone effect (right). Lack of isoplanatism is one of the greatest restrictions to the effectiveness of adaptive optical systems.

As we shall see when we discuss the fundamental photon noise error in a Shack-Hartmann wavefront sensor, in the case of astronomical adaptive optics the "guide star" used to determine the wavefront is unlikely to be the science object under study, whose image is to be improved. The guide star will, in general, be a bright star angularly displaced from the science object, and clearly, from the left hand side of Fig. 5, the wavefront sampled by the guide star is not the same as that from the science object because of this angular displacement. When the wavefronts are, for practical purposes, the same, then we say that the system is "isoplanatic", and when they are not the same we say it is "anisoplanatic". There is an angle Θ_0, called the isoplanatic angle, which quantifies the region over which there is approximate isoplanatism, and it can be shown that this angle is given by:

$$\Theta_0 = \left[2.91 \left(\frac{2\pi}{\lambda}\right)^2 (\cos\gamma)^{-1} \int_0^\infty C_n^2(z) z^{5/3} dz\right]^{-\frac{3}{5}} \qquad \text{rad} \qquad (56)$$

and the mean squared wavefront error σ_Θ^2 for a science object observed Θ away from the guide star is

$$\sigma_\Theta^2 = \left(\frac{\Theta}{\Theta_0}\right)^{5/3} \qquad \text{rad}^2 \qquad (57)$$

Note that like the isoplanatic angle, like the Fried parameter r_0 is proportional to $\lambda^{6/5}$, that is, it is approximately 5.3 times larger (28 times in solid angle) at 2.2μm than at 0.55μm. If we have a single layer of turbulence at height z, then,

$$\Theta_0 \approx 0.31 \frac{r_o}{z} \qquad \text{rad} \qquad (58)$$

Remembering there are $\approx$206265 arcseconds in a radian, then for a typical value of r_0 in the visible of r_0 =10cm and for a layer at altitude z =5km, then the isoplantic angle is only $\approx$1.3 arcseconds — this is a very small angle indeed. This is a somewhat pessimistic estimate (it is too small!) but clearly the isoplanatic angle is very small and this has a profound influence on the effectiveness of adaptive optics.

The isoplanatic angle is also a very important issue for some potential non-astronomical applications of adaptive optics. We have to immediately dispel the notion that large fields of view (many degress) in imaging systems can be corrected using adaptive optics, at least if the distorting medium is extended in three dimensions. If the distortion is all introduced in one plane then there is the possibility of conjugating the correction to that plane.

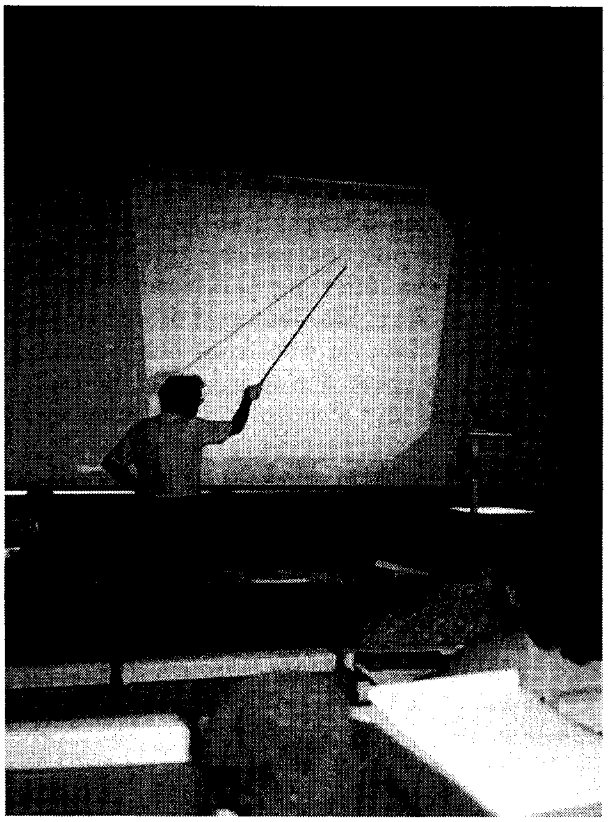
10. Conclusion

In this brief survey, I have highlighted some of the key issues in understanding the optical effects of atmospheric turbulence. For a more thorough understanding, I urge the interested reader to delve into the literature — there is *lots* of it — and to embark on simple measurements and observations. The analysis of practical observations is deeply rewarding, not only for a better understanding of atmospheric effects but also for a real appreciation of random processes in general.

References

1. J-M Conan, Gérard Rousset, and P-Y Madec. Wave-front temporal spectra in high resolution imaging through turbulence. *J Opt Soc Am A*, 12:1559–1570, 1995.
2. D L Fried. Statistics of a geometric representation of wavefront distortion. *J Opt Soc Am*, 55:1427–1435, 1965.
3. D L Fried. Optical resolution through a randomly inhomogeneous medium for very long and very short exposures. *J Opt Soc Am*, 56:1372–1379, 1966.
4. D L Fried. Atmospheric turbulence optical effects: Understanding the adaptive-optics implications. In D M Alloin and J-M Mariotti, editors, *Adaptive Optics in Astronomy*, pages 25–57. Kluwer Academic Publishers, 1994.
5. D L Fried and J F Belsher. Analysis of fundamental limits of artificial guide star adaptive optics system performance for astronomical imaging. *J Opt Soc Am A*, 11:277–287, 1994.
6. D P Greenwood. Bandwidth specifications for adaptive optics. *J Opt Soc Am*, 67:390–392, 1977.
7. R E Hugnagel and N R Stanley. Modulation transfer function with image transmission through turbulent media. *J Opt Soc Am*, 54:52–61, 1964.
8. A N Kolmogorov. The local structure of turbulence in incompressible fluids for very large Reynolds numbers. *Proc Roy Soc Lond A*, 434:9–13, 1991.
9. V N Mahajan. Zernike annular polynomials and optical aberrations of systems with annular pupils. *Opt Phot News*, 5:Engng Lab Notes, 1994.
10. V N Mahajan. Zernike circle polynomials and optical aberrations of systems with circular pupils. *Opt Phot News*, 5:Engng Lab Notes, 1994.
11. F Martin, A Tokovinin, A Ziad, R Conan, J Borgnino, R Avila, A Agabi, and M Sarazan. First statistical data on wavefront outer scale at La Silla Observatory from the GSM instrument. *Astron Astrophys*, 336:L49–L52, 1998.
12. R J Noll. Zernike polynomials and atmospheric turbulence. *J Opt Soc Am*, 66:Suppl 14–1, 307–310, 1975.
13. F Roddier. Effects of atmospheric turbulence in optical astronomy. In E Wolf, editor, *Progress in Optics*, volume 19, chapter 5, pages 281–376. Elsevier, 1981.
14. F Roddier, editor. *Adaptive Optics in Astronomy.* Cambridge University Press, 1999.
15. N Roddier. Atmospheric wavefront simulation using Zernike polynomials. *Opt Engng*, 29:1174–1180, 1990.
16. M Roggemann and B M Welsh. *Imaging Through Turbulence.* CRC Press, 1996.
17. M Sarazin and F Roddier. The ESO differential image motion monitor. *Astron Astrophys*, 227:294–300, 1990.

Coffee break for C. Dainty (left) & E. Lecoarer (right)



R. Foy & the cone effect