Statistical Analysis of Closed-loop AO systems:

An exercise in characteristic functionals

Harrison H. Barrett
University of Arizona

OUTLINE

- Random vectors and characteristic functions
- Random processes and characteristic functionals
- Fully developed speckle
- Residual speckle in closed-loop AO
- Towards a complete statistical analysis of AO images

References

Barrett and Myers, Foundations of Image Science

Chap. 8, Stochastic Properties of Objects and Images

Chap. 11, Poisson Statistics and Photon Counting

Lectures 8 and 13 from last year (on group web site)

Two pending JOSA papers:

Objective Assessment of Image Quality IV

Maximum Likelihood Methods in Wavefront Sensing

Some key points from Lecture 8

An M-dimensional random vector $\mathbf{g}$ is a set of M scalar random variables, $\{g_m, m = 1, ..., M\}$. (Think of measurements from some detector array.)

If each g_m can take a continuous range of values, the vector $\mathbf{g}$ is described by a probability density function (PDF) $\text{pr}(\mathbf{g}) = \text{pr}(\{g_m\})$, and the mean of $\mathbf{g}$ is given by

$$\langle\, \mathbf{g} \,\rangle = \overline{\mathbf{g}} = \{\overline{g}_m\} = \int_{-\infty}^{\infty} dg_1 \int_{-\infty}^{\infty} dg_2 \cdots \int_{-\infty}^{\infty} dg_M \, \mathbf{g} \; \text{pr}(\mathbf{g}) = \int_{\infty} d^M g \; \mathbf{g} \; \text{pr}(\mathbf{g}) \,.$$

If each g_m can take on only integer values, the vector $\mathbf{g}$ is described by a probability (not PDF) $\text{Pr}(\mathbf{g}) = \text{Pr}(\{g_m\})$, and the mean of $\mathbf{g}$ is given by

$$\langle\, \mathbf{g} \,\rangle = \overline{\mathbf{g}} = \{\overline{g}_m\} = \sum_{g_1=0}^{\infty} \sum_{g_2=0}^{\infty} \cdots \sum_{g_M=0}^{\infty} \mathbf{g} \; \text{Pr}(\mathbf{g}) \,.$$

(Note that the sum here is over each of the g_m, not over m.)

Covariance matrix

The *covariance matrix* is a generalization of the variance to random vectors.

For an MD random vector g, the covariance matrix $\mathbf{K}$ is an $M \times M$ matrix with elements given by

$$K_{ij} = \left\langle (g_i - \overline{g}_i)(g_j - \overline{g}_j)^* \right\rangle \ , \tag{8.16}$$

where the asterisk indicates complex conjugate, allowing for the possibility that components of g might be complex. It follows from this definition that $\mathbf{K}$ is Hermitian, *i.e.*, $K_{ij} = K_{ji}^*$.

In outer product form,

$$\mathbf{K} = \left\langle \, [\mathbf{g} - \overline{\mathbf{g}}\,] \; [\mathbf{g} - \overline{\mathbf{g}}\,]^\dagger \, \right\rangle \ .$$

Any random variable covaries with itself. The diagonal elements of the covariance matrix are the variances of the components:

$$K_{jj} = \mathsf{Var}\{g_j\} \, . \tag{8.18}$$

$\boxed{\text{Characteristic functions (from Lecture 8)}}$

For any random variable or vector, the *characteristic function* is the expectation of the appropriate Fourier kernel:

$$\psi(\xi) \equiv \left\langle\, e^{-2\pi i \xi x} \,\right\rangle \, .$$

If x is scalar and real-valued, then

$$\psi(\xi) = \int_{-\infty}^{\infty} dx \; \mathsf{pr}(x)\, e^{-2\pi i \xi x} \, , \tag{C.53}$$

and the PDF and characteristic function form a Fourier transform pair:

$$\mathsf{pr}(x) = \int_{-\infty}^{\infty} d\xi \; \psi(\xi)\, e^{2\pi i \xi x} \, . \tag{C.54}$$

Caution: Do not confuse ξ with a spatial frequency.

Moments of the random variable x can be derived through differentiation of $\psi(\xi)$:

$$\left\langle\, x^k \,\right\rangle = (-2\pi i)^{-k} \, \frac{\partial^k}{\partial \xi^k} \psi(\xi) \bigg|_{\xi=0} \, . \tag{C.55}$$

Characteristic function of a real random vector

For a real $M \times 1$ random vector g (column vector), the characteristic function is defined as

$$\psi_{\mathbf{g}}(\boldsymbol{\xi}) = \left\langle \exp(-2\pi i \boldsymbol{\xi}^t \mathrm{g}) \right\rangle \,, \tag{8.26}$$

where $\boldsymbol{\xi}^t$ is a real $1 \times M$ vector.

For the case of a continuous-valued random vector, $\psi_{\mathbf{g}}(\boldsymbol{\xi})$ can be written as

$$\psi_{\mathbf{g}}(\boldsymbol{\xi}) = \int_\infty d^M g \ \mathrm{pr}(\mathrm{g}) \exp(-2\pi i \boldsymbol{\xi}^t \mathrm{g}) \,. \tag{8.27}$$

This integral is the MD Fourier transform of the PDF, so

$$\mathrm{pr}(\mathrm{g}) = \int_\infty d^M \xi \ \psi_{\mathbf{g}}(\boldsymbol{\xi}) \exp(2\pi i \boldsymbol{\xi}^t \mathrm{g}) \,. \tag{8.28}$$

Linear transformations of random vectors

If the random vector $\mathbf{g}$ is generated as the output of a linear filter acting on the random vector **f**, we can characterize the linear transformation by an $M \times N$ matrix $\mathbf{H}$. Then we can write the $M \times 1$ output vector $\mathbf{g}$ in terms of the $N \times 1$ input vector **f** as

$$\mathbf{g} = \mathbf{H}\mathbf{f}\,. \tag{8.40}$$

From the linearity of the expectation operator, we have immediately for the mean of $\mathbf{g}$,

$$\overline{g} = \langle \mathbf{g} \rangle = \langle \mathbf{H}\mathbf{f}\, \rangle = \mathbf{H}\,\langle \mathbf{f}\, \rangle = \mathbf{H}\overline{\mathbf{f}}\,. \tag{8.49}$$

The covariance matrix of $\mathbf{g}$ is found as

$$\mathbf{K_g} = \left\langle \Delta\mathbf{g}\Delta\mathbf{g}^{\dagger} \right\rangle = \left\langle (\mathbf{H}\mathbf{f} - \mathbf{H}\overline{\mathbf{f}}\,)(\mathbf{H}\mathbf{f} - \mathbf{H}\overline{\mathbf{f}}\,)^{\dagger} \right\rangle = \mathbf{H} \left\langle \Delta\mathbf{f}\,\Delta\mathbf{f}^{\,\dagger} \right\rangle \mathbf{H}^{\dagger} = \mathbf{H}\mathbf{K_f}\mathbf{H}^{\dagger}\,, \tag{8.50}$$

where $\Delta\mathbf{f} \equiv \mathbf{f} - \overline{\mathbf{f}}$.

Digression: Definition of adjoint

Suppose $\boldsymbol{\mathcal{H}}$ is a linear operator mapping space $\mathbb{U}$ to space $\mathbb{V}$, e.g. object space to image space,

$$\mathbf{g}_1 = \boldsymbol{\mathcal{H}}\mathbf{f}_1 \,.$$

Let $\mathbf{g}_2$ be some other vector in $\mathbb{V}$ and compute the scalar product,

$$(\mathbf{g}_2, \mathbf{g}_1)_{\mathbb{V}} = (\mathbf{g}_2, \boldsymbol{\mathcal{H}}\mathbf{f}_1)_{\mathbb{V}} \,.$$

Adjoint operator $\boldsymbol{\mathcal{H}}^\dagger$ is defined by requiring that

$$(\mathbf{g}_2, \boldsymbol{\mathcal{H}}\mathbf{f}_1)_{\mathbb{V}} = (\boldsymbol{\mathcal{H}}^\dagger\mathbf{g}_2, \mathbf{f}_1)_{\mathbb{U}} \,.$$

Alternative notation:

$$\mathbf{g}_2^\dagger\boldsymbol{\mathcal{H}}\mathbf{f}_1 = (\boldsymbol{\mathcal{H}}^\dagger\mathbf{g}_2)^\dagger\mathbf{f}_1$$

Linear transformation of the characteristic function

Transformation of the PDF is tricky, but the characteristic function is easy:

$$\psi_{\mathbf{g}}(\boldsymbol{\xi}) = \Big\langle \exp(-2\pi i \boldsymbol{\xi}^t \mathbf{H}\mathbf{f}\,) \Big\rangle = \Big\langle \exp\Big[-2\pi i (\mathbf{H}^t \boldsymbol{\xi})^t \mathbf{f}\Big] \Big\rangle\,, \tag{8.42}$$

where the last step has used the definition of the adjoint. Thus

$$\psi_{\mathbf{g}}(\boldsymbol{\xi}) = \psi_{\mathbf{f}}(\mathbf{H}^t \boldsymbol{\xi})\,, \tag{8.43}$$

so knowledge of $\psi_{\mathbf{f}}$ and $\mathbf{H}$ immediately gives $\psi_{\mathbf{g}}$.

The PDF on $\mathbf{g}$ can in principle be found by taking an inverse MD Fourier transform of (8.43). Formally, we can write

$$\mathrm{pr}(\mathbf{g}) = \int_{\infty} d^M\xi\ \psi_{\mathbf{f}}(\mathbf{H}^t \boldsymbol{\xi}) \exp(2\pi i \boldsymbol{\xi}^t \mathbf{g})\,, \tag{8.44}$$

but in practice the integral might not be easy. The problem is that we are integrating a function of an ND vector over an MD space.

Intro to random processes

Simple definition: A random process is a random variable that varies with space and/or time. More formally, a spatial random process is a function of two variables, $\mathbf{r}$ and ζ. Depending on the context, $f(\mathbf{r}, \zeta)$ can refer to:

1. The family of spatial functions, referred to as the ensemble; in this case, $\mathbf{r}$ and ζ are variables (e.g., all possible images)

2. A single realization or sample of the spatial functions; in this case, $\mathbf{r}$ is variable and ζ is fixed (one particular image)

3. The random variable at a single point; in this case, $\mathbf{r}$ is fixed and ζ is variable (value at one point in all possible images)

4. A single number; in this case, $\mathbf{r}$ is fixed and ζ is fixed (value at one point in one image)

Notational quirks: Usually drop the ζ designator and hope the differences above will be clear by context

Classifications of random processes

- Spatial vs. temporal (or spatiotemporal)
- Scalar vs. vector valued
- Continuous vs. discrete valued
- Complex or real
- Kind of function:

  Square-integrable (finite "energy")

  Finite power (energy per unit time)

  Generalized (e.g., delta functions)

Averages

Consider a scalar-valued continuous random process. For fixed $\mathbf{r}$, $f(\mathbf{r})$ is simply a random variable (interpretation 3), and its expectation is defined just as for any other random variable;

$$\mathsf{E}\{f(\mathbf{r})\} = \langle f(\mathbf{r})\rangle = \bar{f}(\mathbf{r}) = \int_{-\infty}^{\infty} df(\mathbf{r})\; f(\mathbf{r})\,\mathrm{pr}[f(\mathbf{r})]\,. \tag{8.71}$$

Computation of this expectation requires only the univariate PDF $\mathrm{pr}[f(\mathbf{r})]$.

N.B. The integral is over $f(\mathbf{r})$, not $\mathbf{r}$; result can still be a function of $\mathbf{r}$.

Moments and variance

Moments of $f(\mathbf{r})$ are defined easily. For example, the j^{th} moment is given by

$$\left\langle [f(\mathbf{r})]^j \right\rangle = \int_{-\infty}^{\infty} df(\mathbf{r})\ [f(\mathbf{r})]^j \, \mathrm{pr}[f(\mathbf{r})] \,. \tag{8.72}$$

Again, the integral is over $f(\mathbf{r})$, not over **r**; The resultant, $\left\langle [f(\mathbf{r})]^j \right\rangle$, can still be a function of **r**.

Having defined moments, we can also define the variance of a random process. In the general complex case, the variance is given by

$$\mathrm{Var}\{f(\mathbf{r})\} = \mathsf{E}\left\{ |f(\mathbf{r})| - |\mathsf{E}\{f(\mathbf{r})\}|^2 \right\} = \mathsf{E}\left\{ |f(\mathbf{r})|^2 \right\} - |\mathsf{E}\{f(\mathbf{r})\}|^2$$

$$= \int_{-\infty}^{\infty} df(\mathbf{r})\ |f(\mathbf{r})|^2 \, \mathrm{pr}[f(\mathbf{r})] - \left| \int_{-\infty}^{\infty} df(\mathbf{r})\ f(\mathbf{r}) \, \mathrm{pr}[f(\mathbf{r})] \right|^2 . \tag{8.73}$$

Still need only the univariate density $\mathrm{pr}[f(\mathbf{r})]$

Multiple-point expectations

$$\langle f(\mathbf{r}_1) f(\mathbf{r}_2) \rangle = \int_{-\infty}^{\infty} df(\mathbf{r}_1) \int_{-\infty}^{\infty} df(\mathbf{r}_2) \;\; f(\mathbf{r}_1)\, f(\mathbf{r}_2)\, \mathrm{pr}[f(\mathbf{r}_1), f(\mathbf{r}_2)]\,. \tag{8.74}$$

Here, $f(\mathbf{r}_1)$ and $f(\mathbf{r}_2)$ must be regarded as two *distinct* random variables and pr$[f(\mathbf{r}_1), f(\mathbf{r}_2)]$ is their joint density. Only in very special circumstances will it be possible to write pr$[f(\mathbf{r}_1), f(\mathbf{r}_2)]$ as pr$[f(\mathbf{r}_1)]$ pr$[f(\mathbf{r}_2)]$.

A general two-point moment is defined by

$$\langle [f(\mathbf{r}_1)]^m \, [f(\mathbf{r}_2)]^n \rangle =$$

$$\int_{-\infty}^{\infty} df(\mathbf{r}_1) \int_{-\infty}^{\infty} df(\mathbf{r}_2) \; [f(\mathbf{r}_1)]^m \, [f(\mathbf{r}_2)]^n \, \mathrm{pr}[f(\mathbf{r}_1), f(\mathbf{r}_2)]\,. \tag{8.75}$$

Any moment involving the K points $\mathbf{r}_1, \mathbf{r}_2, ..., \mathbf{r}_K$ can be computed if pr$[f(\mathbf{r}_1), f(\mathbf{r}_2), ..., f(\mathbf{r}_K)]$ is known. If this K-fold joint density is known for all values of each of the $\mathbf{r}_k$, the process is said to be *fully characterized* to order K.

PDFs and characteristic functionals

A spatial random process $f(\mathbf{r})$ is an infinite-dimensional vector in a Hilbert space – *if* every sample function is square-integrable.

Thus, at best, an infinite-dimensional PDF would be needed for a complete characterization: very tricky mathematically.

BUT, an infinite-dimensional characteristic *functional* can always be defined – and often calculated explicitly!

ALL statistical properties of a random process are contained in its characteristic functional.

Characteristic functionals – definition

Recall the definition of the characteristic *function* for a MD real random vector:

$$\psi_{\mathrm{g}}(\boldsymbol{\xi}) = \left\langle \exp(-2\pi i \boldsymbol{\xi}^t \mathrm{g}) \right\rangle , \tag{8.26}$$

Here, $\boldsymbol{\xi}$ is a real $M \times 1$ vector, and $\boldsymbol{\xi}^t \mathrm{g}$ denotes a scalar product.

In the case of a random process $f(\mathbf{r})$, each sample function corresponds to a vector $\mathbf{f}$ in an infinite-dimensional Hilbert space, so the frequency vector $\boldsymbol{\xi}$ in (8.26) must be replaced by an infinite-dimensional vector s in the same Hilbert space as $\mathbf{f}$. That means that s describes a function $s(\mathbf{r})$, so the characteristic function becomes a characteristic *functional* $\Psi_{\mathbf{f}}\{s(\mathbf{r})\}$ or $\Psi_{\mathbf{f}}(\mathrm{s})$ for short. It is defined by

$$\Psi_{\mathbf{f}}(\mathrm{s}) = \langle \exp[-2\pi i (\mathrm{s}, \mathbf{f})] \rangle , \tag{8.94}$$

where $(\mathrm{s}, \mathbf{f})$ is the usual $\mathbb{L}_2$ scalar product.

Propagation of characteristic functionals

Consider a linear continuous-to-discrete mapping without noise: $\mathbf{g} = \boldsymbol{\mathcal{H}}\mathbf{f}$, where $\mathbf{f}$ is a function and $\mathbf{g}$ is an $M \times 1$ vector.

Characteristic *functional* of $\mathbf{f}$ is defined by

$$\Psi_{\mathbf{f}}(\mathbf{s}) = \langle \exp[-2\pi i(\mathbf{s}, \mathbf{f})] \rangle \,, \tag{8.94}$$

Characteristic *function* of $\mathbf{g}$ is given by

$$\psi_{\mathbf{g}}(\boldsymbol{\xi}) = \langle \exp[-2\pi i(\boldsymbol{\xi}, \boldsymbol{\mathcal{H}}\mathbf{f})] \rangle = \langle \exp[-2\pi i(\boldsymbol{\mathcal{H}}^{\dagger}\boldsymbol{\xi}, \mathbf{f})] \rangle \,, \tag{8.95}$$

where the second step follows from the definition of the adjoint,

Comparison of (8.94) and (8.95) shows that

$$\psi_{\mathbf{g}}(\boldsymbol{\xi}) = \Psi_{\mathbf{f}}(\boldsymbol{\mathcal{H}}^{\dagger}\boldsymbol{\xi}) \,, \tag{8.96}$$

Effect of measurement noise

Now assume $\mathbf{g} = \boldsymbol{\mathcal{H}}\mathbf{f} + \mathbf{n}$, where $\mathbf{n}$ is an $M \times 1$ random vector.

Characteristic function of $\mathbf{g}$:

$$\psi_{\mathbf{g}}(\boldsymbol{\xi}) = \langle \exp[-2\pi i(\boldsymbol{\xi}^\dagger \boldsymbol{\mathcal{H}}\mathbf{f} + \boldsymbol{\xi}^\dagger \mathbf{n})] \rangle_{\mathbf{n},\mathbf{f}} = \langle \langle \exp[-2\pi i(\boldsymbol{\xi}^\dagger \boldsymbol{\mathcal{H}}\mathbf{f} + \boldsymbol{\xi}^\dagger \mathbf{n})] \rangle_{\mathbf{n}|\mathbf{f}} \rangle_{\mathbf{f}} .$$

No loss of generality to this point, but if $\mathbf{n}$ is independent of $\mathbf{f}$, then

$$\psi_{\mathbf{g}}(\boldsymbol{\xi}) = \psi_{\mathbf{n}}(\boldsymbol{\xi})\, \Psi_{\mathbf{f}}(\boldsymbol{\mathcal{H}}^\dagger \boldsymbol{\xi}) .$$

No surprise – adding independent RVs $\Rightarrow$ convolution of PDFs $\Rightarrow$ multiplication in Fourier domain. Tricky part is that "Fourier domain" for $\mathbf{f}$ is infinite dimensional, PDF not readily defined.

Propagation of characteristic functional with Poisson noise
Clarkson et al., Optics Express, 2002

If $\Pr(\mathbf{g}|\mathbf{f})$ is Poisson, previous result is modified to

$$\psi_{\mathbf{g}}(\boldsymbol{\xi}) = \Psi_{\mathbf{f}}[\boldsymbol{\mathcal{H}}^{\dagger}\,\boldsymbol{\Gamma}(\boldsymbol{\xi})]\,, \tag{8.339}$$

where $\boldsymbol{\Gamma}$ is an operator that acts independently on each component of its vector operand; it is defined such that

$$[\boldsymbol{\Gamma}(\boldsymbol{\xi})]_m = \frac{-1 + \exp(-2\pi i\,\xi_m)}{-2\pi i}\,. \tag{8.338}$$

Summary so far

- Characteristic functionals can be defined for any square-integrable random process (and computed analytically in many cases)

- It is easy to propagate characteristic functionals through linear transformations

- In particular, characteristic functionals for random objects can be propagated through linear imaging systems

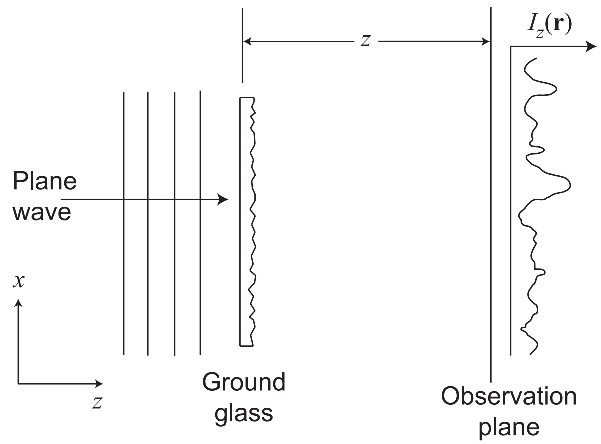
- Photon noise and readout noise in the detector are also easily incorporated

Now let's apply these tools to speckle

- Fully developed Gaussian speckle in free-space propagation
- Fully developed Gaussian speckle in coherent imaging
- Weak speckle in AO

Speckle in a nonimaging system (free-space propagation)

Consider ground glass in plane $z = 0$ illuminated by a monochromatic plane wave propagating in z direction. Observe field and irradiance in a parallel plane a distance z away.



Amplitude transmittance of ground glass:

$$t_{GG}(\mathbf{r}) = \exp[i\phi(\mathbf{r})]\, S(\mathbf{r})\,, \tag{18.9}$$

where $S(\mathbf{r})$ is a binary support function.

Assume the ground glass completely randomizes the phases. In terms of the univariate or single-point PDF for the phase:

$$\text{pr}[\phi(\mathbf{r})] = \frac{1}{2\pi}\,\text{rect}\left(\frac{\phi(\mathbf{r})}{2\pi}\right) \qquad \text{(for all } \mathbf{r}\text{)}\,. \tag{18.11}$$

It follows that

$$\langle t_{GG}(\mathbf{r})\rangle = 0\,, \tag{18.12}$$

where the average is over an ensemble of ground glasses.

Note carefully that $t_{GG}(\mathbf{r})$ is not Gaussian since $|t_{GG}(\mathbf{r})| = 1$

Field in the observation plane

Field emerging from the ground glass:

$$u_0(\mathbf{r}) = e^{ikz}|_{z=0}\, t_{GG}(\mathbf{r}) = t_{GG}(\mathbf{r})\,. \tag{18.15}$$

Propagation to the observation plane:

$$\mathbf{u}_z = \boldsymbol{\mathcal{P}}_z \mathbf{u}_0\,, \tag{18.16}$$

where $\boldsymbol{\mathcal{P}}_z$ is a propagation operator. Kernel of propagator is spherical wave (Huygens wavelet). In the Fresnel approximation,

$$u_z(\mathbf{r}) = \frac{\exp(ikz)}{i\lambda z} \int_\infty d^2 r_0\; u_0(\mathbf{r}_0) \exp\left(i\pi \frac{|\mathbf{r} - \mathbf{r}_0|^2}{\lambda z} \right)\,. \tag{18.17}$$

Like a shift-invariant imaging system where the PSF is a spherical wave or quadratic phase factor. Key difference: *All* points on object contribute to field at *each* point in observation plane.

Propagation of characteristic functionals

The field emerging from the ground glass in the plane $z = 0$ has a characteristic functional given by

$$\Psi_{\mathbf{u}_0}(\boldsymbol{\xi}) = \left\langle \exp[-i\pi(\boldsymbol{\xi}^\dagger \mathbf{u}_0 + \mathbf{u}_0{}^\dagger \boldsymbol{\xi})] \right\rangle . \tag{18.19}$$

The characteristic functional for the propagated field is given by

$$\Psi_{\mathbf{u}_z}(\boldsymbol{\xi}) = \left\langle \exp[-i\pi(\boldsymbol{\xi}^\dagger \mathbf{u}_z + \mathbf{u}_z{}^\dagger \boldsymbol{\xi})] \right\rangle$$

$$= \left\langle \exp\left\{ -i\pi[\boldsymbol{\xi}^\dagger(\boldsymbol{\mathcal{P}}_z \mathbf{u}_0) + (\boldsymbol{\mathcal{P}}_z \mathbf{u}_0)^\dagger \boldsymbol{\xi})] \right\} \right\rangle = \Psi_{\mathbf{u}_0}(\boldsymbol{\mathcal{P}}_z^\dagger \boldsymbol{\xi}) . \tag{18.20}$$

Invocation of central-limit theorem

For derivation of central-limit theorem for complex random processes, see Sec. 18.2.4 in B&M.

With this theorem, can write at once that

$$\Psi_{\mathbf{u_Z}}(\boldsymbol{\xi}) = \mathrm{exp}\left(-\pi^2 \boldsymbol{\xi}^\dagger \boldsymbol{\mathcal{K}}_{\mathbf{u_Z}} \boldsymbol{\xi}\right) , \tag{18.21}$$

where $\boldsymbol{\mathcal{K}}_{\mathbf{u_Z}}$ is the autocovariance operator. To be explicit, the quadratic form is given by

$$\boldsymbol{\xi}^\dagger \boldsymbol{\mathcal{K}}_{\mathbf{u_Z}} \boldsymbol{\xi} = \int_\infty d^2 r \int_\infty d^2 r' \, \xi^*(\mathbf{r}) \, K_{\mathbf{u_Z}}(\mathbf{r}, \mathbf{r}') \, \xi(\mathbf{r}') \, . \tag{18.22}$$

Since $\boldsymbol{\mathcal{K}}_{\mathbf{u_Z}}$ is positive-definite, $\boldsymbol{\xi}^\dagger \boldsymbol{\mathcal{K}}_{\mathbf{u_Z}} \boldsymbol{\xi}$ is real.

Punchline: If we know the autocovariance function, we know all statistics of the field in the observation plane

Propagation of autocovariance function

Assume that the correlation length of the ground glass is short (as it must be for a diffuse scatterer):

$$K_{GG}(\mathbf{r}, \mathbf{r}') \approx S(\mathbf{r})\, \ell_c^2\, \delta(\mathbf{r} - \mathbf{r}')\,, \tag{18.13}$$

Transform this operator with propagator to get:

$$\boldsymbol{\mathcal{K}}_{\mathbf{u}\mathsf{z}} = \boldsymbol{\mathcal{P}}_z \boldsymbol{\mathcal{K}}_{\mathbf{u}_0} \boldsymbol{\mathcal{P}}_z^{\dagger} \approx \ell_c^2\, \boldsymbol{\mathcal{P}}_z \boldsymbol{\mathcal{P}}_z^{\dagger}\,. \tag{18.23}$$

In the Fresnel approximation, the kernel of the operator $\boldsymbol{\mathcal{K}}_{\mathbf{u}\mathsf{z}}$ is

$$\begin{aligned} K_{\mathbf{u}\mathsf{z}}(\mathbf{r}, \mathbf{r}') &= \frac{\ell_c^2}{\lambda^2 z^2} \int_A d^2 r'' \, \mathsf{exp}\left(i \frac{\pi}{\lambda z} |\mathbf{r} - \mathbf{r}''|^2 \right) \mathsf{exp}\left(-i \frac{\pi}{\lambda z} |\mathbf{r}' - \mathbf{r}''|^2 \right) \\ &= \frac{\ell_c^2}{\lambda^2 z^2} L^2\, \mathsf{sinc}\left(\frac{L}{\lambda z} |\mathbf{r} - \mathbf{r}'| \right)\,, \end{aligned} \tag{18.24}$$

Thus the scale of the correlations in the observation plane is determined by the angular subtense of the ground glass, L/z.

Statistics of the irradiance

Irradiance on observation plane is a real random process given by

$$I_z(\mathbf{r}) = |u_z(\mathbf{r})|^2$$

Since this mapping is nonlinear, cannot simply propagate the characteristic functional.

[Many pages of math]

$$\Psi_{\mathbf{I}\ \mathrm{im}}(\zeta) = \frac{1}{\det(\mathcal{I} + 2\pi i\, \mathcal{K}_{\mathbf{u}_{\mathrm{im}}} \mathcal{Z})}\,, \tag{18.102}$$

where $\mathcal{Z}$ is the integral operator with kernel $\zeta(\mathbf{r})\,\delta(\mathbf{r}-\mathbf{r}')$, and the determinant of an integral operator is interpreted as the product of its eigenvalues.

From this characteristic functional we can derive the univariate or single-point characteristic function simply by setting

$$\zeta(\mathbf{r}) = \nu\delta(\mathbf{r} - \mathbf{r}_1)\,.$$

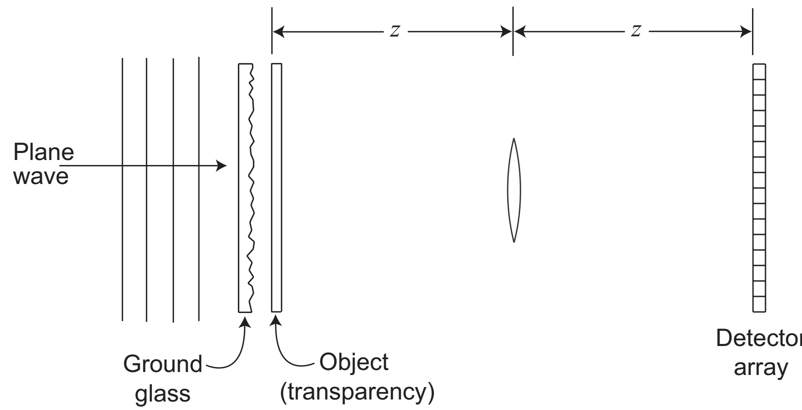
Can show that

$$\psi_{I_{\mathsf{im}}(\mathbf{r}_1)}(\nu) = \frac{1}{1 + 2\pi i\nu\, K_{\mathbf{u}_{\mathsf{im}}}(\mathbf{r}_1, \mathbf{r}_1)}\,, \tag{18.103}$$

from which it follows that $I_{im}(\mathbf{r}_1)$ is a chi-squared random variable with two degrees of freedom, or equivalently an exponentially-distributed random variable.

The mean irradiance at point $\mathbf{r}_1$ is $K_{\mathbf{u}_{\mathsf{im}}}(\mathbf{r}_1, \mathbf{r}_1)$, or simply $\langle |u_{im}(\mathbf{r}_1)|^2\rangle$.

Speckle in an imaging system



Object: Photographic transparency with amplitude transmittance t_{obj}, placed directly over a ground glass in plane $z = 0$

Illumination: Plane wave in z direction

Imaging element: Thin lens of focal length f placed a distance $2f$ from the object.

Detector: Discrete array placed in the focal plane, a distance $2f$ from lens.

Coherent imaging equations

$$u_{im}(\mathbf{r}) = \int_\infty d^2 r_0 \, p_{coh}(\mathbf{r} - \mathbf{r}_0; \mathbf{r}_0) \, u_0(\mathbf{r}_0) \,, \qquad (18.86)$$

where $u_{im}(\mathbf{r})$ is the field in the image plane, $u_0(\mathbf{r}_0)$ is the field emerging from the object transparency in the plane $z = 0$, and $p_{coh}(\mathbf{r} - \mathbf{r}_0; \mathbf{r}_0)$ is the coherent point spread function (PSF).

If there are no field-dependent aberrations (coma and astigmatism), $p_{coh}(\mathbf{r}; \mathbf{r}_0)$ is independent of the second variable and (18.86) is a convolution.

In abstract form (with or without aberrations)

$$\mathbf{u}_{im} = \boldsymbol{\mathcal{P}}_{coh} \mathbf{u}_0 \,. \qquad (18.88)$$

Irradiance and CD mapping by the detector

Convert image-plane field to irradiance:

$$I_{im}(\mathbf{r}) = |u_{im}(\mathbf{r})|^2 \,. \tag{18.89}$$

Consider CD mapping by the detector array. Assume detector output is linear in irradiance:

$$\overline{g}_m = \int_\infty d^2r \, w_m(\mathbf{r}) \, I_{im}(\mathbf{r}) \,, \tag{18.90}$$

where $w_m(\mathbf{r})$ is response function for the m^{th} detector.

Object transparency

$$u_0(\mathbf{r}) = u_{GG}(\mathbf{r})\, t_{obj}(\mathbf{r})\,, \tag{18.91}$$

where $t_{obj}(\mathbf{r})$ is the complex amplitude transmission of the object, and $u_{GG}(\mathbf{r})$ is the field emerging from the ground glass and incident on the object.

Abstractly,

$$\mathbf{u}_0 = \mathbf{u}_{GG}\, \mathbf{t}_{obj}\,. \tag{18.92}$$

The CC propagation rule, (18.88), now becomes

$$\mathbf{u}_{im} = \boldsymbol{\mathcal{P}}_{coh}\{\mathbf{u}_{GG}\, \mathbf{t}_{obj}\}\,. \tag{18.93}$$

Propagation of characteristic functionals

Characteristic functional for image-plane field

$$\Psi_{\mathbf{u}_{\mathsf{im}}}(\boldsymbol{\xi}) = \Psi_{\mathbf{u}_0}(\boldsymbol{\mathcal{P}}^{\dagger}_{coh}\boldsymbol{\xi}) = \Psi_{\mathbf{u}_{GG}}(\mathbf{t}^{*}_{obj}\boldsymbol{\mathcal{P}}^{\dagger}_{coh}\,\boldsymbol{\xi})\,. \tag{18.96}$$

Invoke central limit theorem, assume correlation length of ground glass is short compared to system resolution. Get circular Gaussian:

$$\Psi_{\mathbf{u}_{\mathsf{im}}}(\boldsymbol{\xi}) = \exp(-\pi^2\boldsymbol{\xi}^{\dagger}\boldsymbol{\mathcal{K}}_{\mathbf{u}_{\mathsf{im}}}\boldsymbol{\xi})\,. \tag{18.97}$$

where

$$K_{\mathbf{u}_{\mathsf{im}}}(\mathbf{r},\mathbf{r}') = \ell_c^2 \int d^2r_0\, p_{coh}(\mathbf{r}-\mathbf{r}_0;\mathbf{r}_0)\, p^{*}_{coh}(\mathbf{r}'-\mathbf{r}_0;\mathbf{r}_0)\,|t_{obj}(\mathbf{r}_0)|^2\,. \tag{18.100}$$

Noise is not stationary even if system is shift-invariant.

Irradiance in the image plane

Mean irradiance in image plane given by

$$\langle I_{im}(\mathbf{r})\rangle = \left\langle |u_{im}(\mathbf{r})|^2 \right\rangle = K_{\mathbf{u}_{\mathsf{im}}}(\mathbf{r},\mathbf{r})$$

$$= \ell_c^2 \int d^2r_0 \, |p_{coh}(\mathbf{r}-\mathbf{r}_0;\mathbf{r}_0)|^2 \, |t_{obj}(\mathbf{r}_0)|^2 \, . \tag{18.101}$$

Compare to incoherent imaging.

Characteristic functional (with circular-Gaussian assumption):

$$\Psi_{\mathbf{I}_{\mathsf{im}}}(\zeta) = \frac{1}{\mathrm{det}(\mathcal{I} + 2\pi i \, \mathcal{K}_{\mathbf{u}_{\mathsf{im}}} \mathcal{Z})} \, , \tag{18.102}$$

where $\mathcal{Z}$ is the integral operator with kernel $\zeta(\mathbf{r})\,\delta(\mathbf{r}-\mathbf{r}')$.

Neat point: All statistics of *irradiance* are determined solely by autocorrelation function of *field* (in which the coherent PSF is hidden).

Residual speckle in closed-loop AO

Why is AO different?

- Random phase is in pupil, not object
- Atmospheric phase perturbations have long-range correlations
- Residual phase variations in closed-loop AO are small
- Cannot use theory of fully developed speckle
- Cannot treat fields as circular Gaussian
- Now have three sources of randomness: Object, PSF, image

Basic formulation

General approach: look for characteristic functional for pupil *phase*, not field.

Treat pupil phase as spatial random process (ignore temporal variation of atmosphere and loop dynamics for now)

The instantaneous phase in the pupil has the form,

$$\phi(\mathbf{r}) = \phi_{atm}(\mathbf{r}) - \sum_{n=1}^{N} \alpha_n \, u_n(\mathbf{r}) \,, \tag{1}$$

where $\{u_n(\mathbf{r}), n = 1, ..., N\}$ is an orthonormal basis for DM space and $\{\alpha_n, n = 1, ..., N\}$ is the associated set of control signals expressed in the orthonormal basis.

Let $\boldsymbol{\alpha}$ be an $N \times 1$ vector of coefficients and $\mathbf{u}$ be $N \times 1$ vector of modal functions

$\boxed{\text{Characteristic functional for the pupil phase}}$

Definition:

$$\Psi_{\phi}(\mathrm{s}) \equiv \left\langle \exp(-2\pi i \mathrm{s}^{\dagger}\boldsymbol{\phi}) \right\rangle$$

With (1),

$$\Psi_{\phi}(\mathrm{s}) = \left\langle \exp\left[-2\pi i \mathrm{s}^{\dagger}\left(\boldsymbol{\phi}_{atm} - \sum_{n=1}^{N} \alpha_n \, \mathbf{u}_n\right)\right] \right\rangle$$

$$= \left\langle \left\langle \exp\left[-2\pi i \, \mathrm{s}^{\dagger}\left(\boldsymbol{\phi}_{atm} - \sum_{n=1}^{N} \alpha_n \, \mathbf{u}_n\right)\right] \right\rangle_{\boldsymbol{\alpha}|\boldsymbol{\phi}_{atm}} \right\rangle_{\boldsymbol{\phi}_{atm}}$$

$$= \left\langle \exp\left(-2\pi i \, \mathrm{s}^{\dagger}\boldsymbol{\phi}_{atm}\right) \left\langle \exp\left(2\pi i \sum_{n=1}^{N} \alpha_n \, \mathbf{u}_n^{\dagger}\mathbf{s}\right) \right\rangle_{\boldsymbol{\alpha}|\boldsymbol{\phi}_{atm}} \right\rangle_{\boldsymbol{\phi}_{atm}} . \quad (2)$$

From last slide:

$$\Psi_{\phi}(\mathrm{s}) = \left\langle \exp\left(-2\pi i\, \mathrm{s}^{\dagger} \boldsymbol{\phi}_{atm}\right) \left\langle \exp\left(2\pi i \sum_{n=1}^{N} \alpha_n\, \mathbf{u}_n^{\dagger}\mathbf{s}\right) \right\rangle_{\boldsymbol{\alpha}|\boldsymbol{\phi}_{atm}} \right\rangle_{\boldsymbol{\phi}_{atm}} .$$

Conditional characteristic *function* for $\boldsymbol{\alpha}$ is defined by

$$\psi_{\boldsymbol{\alpha}|\boldsymbol{\phi}_{atm}}(\boldsymbol{\xi}) \equiv \left\langle \exp\left(-2\pi i\, \boldsymbol{\xi}^{\dagger}\boldsymbol{\alpha}\right) \right\rangle_{\boldsymbol{\alpha}|\boldsymbol{\phi}_{atm}} = \left\langle \exp\left(-2\pi i \sum_{n=1}^{N} \xi_n\, \alpha_n\right) \right\rangle_{\boldsymbol{\alpha}|\boldsymbol{\phi}_{atm}} ,$$

where $\boldsymbol{\xi}$ is an $N \times 1$ real vector.

Thus,

$$\Psi_{\phi}(\mathrm{s}) = \left\langle \exp\left(-2\pi i\, \mathrm{s}^{\dagger} \boldsymbol{\phi}_{atm}\right) \psi_{\boldsymbol{\alpha}|\boldsymbol{\phi}_{atm}}(-\,\mathbf{u}^{\dagger}\mathrm{s}) \right\rangle_{\boldsymbol{\phi}_{atm}} . \qquad (3)$$

No approximations so far.

Assumptions about the wavefront sensor

- Estimates of modal coefficients are unbiased:

$$\overline{\boldsymbol{\alpha}}_n \equiv \langle\, \boldsymbol{\alpha}_n \,\rangle_{\boldsymbol{\alpha}|\phi_{atm}} = \mathbf{u}_n^{\dagger}\boldsymbol{\phi}_{atm}$$

- Estimates are jointly normal (possibly correlated):

$$\psi_{\boldsymbol{\alpha}|\phi_{atm}}(\boldsymbol{\xi}) = \exp\left[-2\pi i \boldsymbol{\xi}^{\dagger}\left(\mathbf{u}^{\dagger}\boldsymbol{\phi}_{atm}\right)\right] \exp\left[-2\pi^2 \boldsymbol{\xi}^{\dagger}\mathbf{K}_{\boldsymbol{\alpha}|\phi_{atm}}\,\boldsymbol{\xi}\right]$$

- Covariance matrix for coefficients does not depend on ϕ_{atm}. Valid in any of the following cases:

  – Wavefront sensor is dominated by readout noise;

  – Wavefront sensor is dominated by Poisson noise from diffuse background;

  – Wavefront presented to the WFS in closed loop is nearly flat

THE RESULT

$$\Psi_{\phi}(\mathrm{s}) = \mathsf{exp}\left[-2\pi^2\,\mathbf{s}^{\dagger}\mathbf{u}\mathbf{K}_{\boldsymbol{\alpha}}\mathbf{u}^{\dagger}\mathbf{s}\right]\Psi_{\phi_{atm}}(\boldsymbol{\mathcal{P}}_{\perp}\mathbf{s})\ . \tag{4}$$

where $\boldsymbol{\mathcal{P}}_{\perp}$ is the projection operator onto the orthogonal complement of DM space.

Thus the second factor represents the atmospheric modes that cannot be corrected by the specific deformable mirror.

The first factor represents the noise in the wavefront sensor propagated into the pupil phase.

This characteristic function contains all of the statistic properties of an AO system, yet it is something we can write down analytically and evaluate in many practical cases of interest.

Kolmogorov statistics

For pure Kolmogorov turbulence, $\phi_{atm}(\mathbf{r})$ is a zero-mean, real-valued Gaussian random process, so

$$\Psi_{\phi_{atm}}(\mathbf{s}) = \exp\left[-2\pi^2 \mathbf{s}^\dagger \mathcal{K}_{atm}\, \mathbf{s}\right] \ , \tag{5}$$

where $\mathcal{K}_{atm}$ is the autocovariance operator, with kernel defined by

$$\left[\mathcal{K}_{atm}\right](\mathbf{r}, \mathbf{r}') \equiv \left\langle\, \phi(\mathbf{r})\, \phi(\mathbf{r}')\, \right\rangle$$

Note that $\left[\mathcal{K}_{atm}\right](\mathbf{r}, \mathbf{r}')$ is almost but not quite a wide-sense stationary autocovariance function; it is a function of $\mathbf{r} - \mathbf{r}'$ if both points are within the pupil, but the pupil support spoils the stationarity.

With (5), an immediate consequence of (4) is that the corrected phase is also a zero-mean Gaussian random process.

Departures from Kolmogorov

Suppose r_0 is random. Then the atmospheric characteristic functional becomes

$$\Psi_{\phi_{atm}}(\mathrm{s}) = \left\langle \mathsf{exp}\left\{-2\pi^2 \mathrm{s}^\dagger [\mathcal{K}_{atm}(r_0)]\,\mathrm{s}\right\}\right\rangle_{r_0}$$

$$= \int_0^\infty dr_0 \ \mathsf{pr}(r_0) \ \mathsf{exp}\left\{-2\pi^2 \mathrm{s}^\dagger [\mathcal{K}_{atm}(r_0)]\,\mathrm{s}\right\} \ . \qquad (6)$$

Now the uncorrected phase is no longer a Gaussian random process, but it is possible that the corrected phase will be Gaussian. If the density of actuators is high and $\boldsymbol{\mathcal{P}}_\perp \mathrm{s} \approx \mathbf{0}$ for all functions s of interest, then the corrected phase might be dominated by the first factor in (4).

Incoherent PSF and pinned speckle

Suppose there is only a single on-axis star present in the FOV of the telescope and light is narrowband with mean wavelength λ.

The complex scalar field at point $\mathbf{r}_d$ in the image plane is given by

$$u(\mathbf{r}_d) \propto \int_\infty d^2r\, a_{ap}(\mathbf{r})\, \mathrm{exp}[i\phi(\mathbf{r})]\ \mathrm{exp}\left(\frac{2\pi i \mathbf{r}\cdot\mathbf{r}_d}{\lambda f}\right)\ .$$

For weak phase,

$$\mathrm{exp}[i\phi(\mathbf{r})] = 1 + i\,\phi(\mathbf{r}) - \tfrac{1}{2}[\phi(\mathbf{r})]^2 + \cdots$$

Random image-plane field:

$$u(\mathbf{r}_d) \propto A_{ap}(\tilde{\mathbf{r}}) + i[A_{ap} * \Phi](\tilde{\mathbf{r}}) - \tfrac{1}{2}[A_{ap} * \Phi * \Phi](\tilde{\mathbf{r}}) + \cdots, \qquad (7)$$

where $A_{ap}(\boldsymbol{\rho})$ and $\Phi(\boldsymbol{\rho})$ are the 2D spatial Fourier transforms of $a_{ap}(\mathbf{r})$ and $\phi(\mathbf{r})$, respectively, and $\tilde{\mathbf{r}} \equiv \mathbf{r}_d/(\lambda f)$.

Incoherent PSF and pinned speckle – cont.

From last slide, random image-plane field:

$$u(\mathbf{r}_d) \propto A_{ap}(\tilde{\mathbf{r}}) + i[A_{ap} * \Phi](\tilde{\mathbf{r}}) - \tfrac{1}{2}[A_{ap} * \Phi * \Phi](\tilde{\mathbf{r}}) + \cdots, \qquad (7)$$

Random image-plane irradiance:

$$I(\mathbf{r}_d) = |u(\mathbf{r}_d)|^2 \propto |A_{ap}(\tilde{\mathbf{r}})|^2 - 2\,A_{ap}(\tilde{\mathbf{r}})\,\mathrm{Im}\left\{[A_{ap} * \Phi](\tilde{\mathbf{r}})\right\}$$

$$+|[A_{ap} * \Phi](\tilde{\mathbf{r}})|^2 - A_{ap}(\tilde{\mathbf{r}})\,\mathrm{Re}\left\{[A_{ap} * \Phi * \Phi](\tilde{\mathbf{r}})\right\} + \cdots, \qquad (8)$$

Conclusions:

- Field in focal plane is not Gaussian
- Linear term in field expansion is Gaussian but not circular Gaussian
- Linear term in irradiance *is* Gaussian (Gaussian $\phi \Rightarrow$ Gaussian Φ)
- Quadratic and higher terms in irradiance are not Gaussian
- Residual speckle pattern is modulated (pinned) by Airy pattern
- Need a characteristic functional for FT of residual pupil phase!

$\boxed{\text{Characteristic functional for } \Phi(\boldsymbol{\rho})}$

Definition:

$$\Psi_{\mathbf{\Phi}}(\mathrm{S}) \equiv \left\langle \exp\left(-2\pi i\, \mathbf{S}^{\dagger}\mathbf{\Phi}\right)\right\rangle\ ,$$

where $\mathbf{S} \Rightarrow$ an arbitrary frequency-domain function $S(\boldsymbol{\rho})$.

It follows from Parseval's theorem that $\mathbf{S}^{\dagger}\mathbf{\Phi} = \mathrm{s}^{\dagger}\boldsymbol{\phi}$ so we have

$$\Psi_{\mathbf{\Phi}}(\mathrm{S}) = \left\langle \exp\left(-2\pi i\, \mathrm{s}^{\dagger}\boldsymbol{\phi}\right)\right\rangle = \Psi_{\boldsymbol{\phi}}(\boldsymbol{\mathcal{F}}^{-1}\mathrm{S})\,. \tag{9}$$

This simple result, a special case of our transformation rule, takes advantage of the Fourier transform being unitary.

Why we *don't* need a characteristic functional for the irradiance

All statistics of the irradiance can be obtained from (8) and (9), even without an explicit expression for the characteristic functional of the irradiance. The trick is to make judicious choices for the function $S(\boldsymbol{\rho})$ in the argument of $\Psi_{\boldsymbol{\Phi}}(\mathrm{S})$.

Example: Suppose we want to evaluate moments of a product of the form $\Phi(\boldsymbol{\rho}_1)\,\Phi(\boldsymbol{\rho}_2)$. For this purpose, we need a characteristic *function* defined by

$$\psi_{\Phi(\boldsymbol{\rho}_1),\Phi(\boldsymbol{\rho}_2)}(\nu_1,\nu_2) \equiv \langle\ \exp\{-2\pi i\,[\nu_1\,\Phi(\boldsymbol{\rho}_1)+\nu_2\,\Phi(\boldsymbol{\rho}_2)\,]\}\ \rangle\,.$$

Now suppose we set

$$S(\boldsymbol{\rho}) = \nu_1\,\delta(\boldsymbol{\rho}-\boldsymbol{\rho}_1)+\nu_2\,\delta(\boldsymbol{\rho}-\boldsymbol{\rho}_2)\,.$$

From (9) and the sifting property of the delta function:

$$\begin{aligned}\psi_{\Phi(\boldsymbol{\rho}_1),\Phi(\boldsymbol{\rho}_2)}(\nu_1,\nu_2) &= \Psi_{\boldsymbol{\Phi}}[\nu_1\,\delta(\boldsymbol{\rho}-\boldsymbol{\rho}_1)+\nu_2\,\delta(\boldsymbol{\rho}-\boldsymbol{\rho}_2)] \\ &= \Psi_{\phi}[\nu_1\ \exp(2\pi i(\boldsymbol{\rho}_1\cdot\mathbf{r})+\nu_2\ \exp(2\pi i\boldsymbol{\rho}_2\cdot\mathbf{r})]\,.\end{aligned}$$

Why we *do* need a characteristic functional for the irradiance

Short answer: so we can propagate it through a discrete detector array and add noise.

$\boxed{\text{Still to do}}$

- Try to find characteristic functional for image irradiance
- Find M-dimensional characteristic function for data $\mathbf{g}$
- Incorporate time dependence of atmosphere and control loop
- Find MJ-dimensional characteristic function for data sequence $\mathbf{G}$ (M pixels in each of J image frames)
- Relate all of the above to observer performance