1 Intro

The goal of this note is to A) write down how I think the HuygensPSF tool in Zemax works, (which is also the description of how it works in my raytracing software BATOID), and B) connect that HuygensPSF description to Fourier optics, especially the relationships between the various 2D grids of points (lattices) that arise in Fourier optics.

2 HuygensPSF

The Huygens construction electric field amplitude given a set of monochromatic plane waves K evaluated at a particular time t and location $\vec{x}$ is:

$$A(\vec{x},t) \propto \sum_{\vec{k}\in K} \exp\left(i\phi_{\vec{k}}\right) \exp\left(i\vec{k}\cdot\vec{x} - \omega t\right) \tag{1}$$

In the above, I use $\vec{k}$ both as an index over different plane waves and to indicate the wave vector of an individual plane wave. The phase of each plane wave is given by $\phi_{\vec{k}}$, and the temporal frequency of each wave is ω.

Note that the input to the sum and therefore the resultant amplitude are both scalars. I.e., I'm asserting a scalar diffraction theory. A more accurate vector diffraction theory would replace the $\exp\left(\phi_{\vec{k}}\right)$ with something like $\vec{E}_{\vec{k}} \exp\left(i\phi_{\vec{k}}\right)$ and the resulting amplitude would also be a vector quantity.

The PSF intensity given by Eq. 1 is then given by the square of the electric field amplitude, which in our scalar field theory is:

$$I(\vec{x}) = I(\vec{x},t) = |A(\vec{x},t)|^2 \tag{2}$$

Note that the time dependence drops out in the amplitude-squared operation. Given the above, the whole trick then is to figure out what is the set K. We will assert here that K can be developed by tracing a sufficiently dense initial set of rays (all representing the same initial incident plane wave) through the optical system in question and then reinterpretting each ray as a plane wave as it impacts the focal surface.

3 Lattices

I will use the word "lattice" to refer to a two dimensional set of points that are uniformly distributed in the plane on a grid, which may be square or rectilinear, but is not required to be. (I.e., a Bravais lattice as borrowed from solid-state physics.) Concretely, the lattice L is the set of points:

$$L = \left\{ \vec{\ell}_1 \mathrm{i} + \vec{\ell}_2 \mathrm{j} \;\middle|\; \mathrm{i},\mathrm{j} \in \mathbb{Z}\times\mathbb{Z} \right\} \tag{3}$$

where $\vec{\ell}_1$ and $\vec{\ell}_2$ are the "primitive lattice vectors", whose only constraint, I believe, is that they are linearly independent. In particular, they need not be unit vectors or orthogonal to each other. (Note also the use of romanized i and j as indices, distinct from the italicized i used for the imaginary unit.)

In what follows, we will be interested in at least 5 different lattices, itemized below together with their respective dimensions:

- The pupil lattice U [length]
- The focal plane lattice X [length]

- The Fourier conjugate of the focal plane lattice K_{dft} [inverse length]
- The sky tangent plane lattice Θ [angle]
- The Fourier conjugate of the sky tangent plane Q [inverse angle].

While the specific point grids of interest on the focal plane or on the sky won't generally include the origin, for this note we'll ignore such origin shifts. In fact, for Fourier optics we are generally interested in just the lattice subsets for which $\mathrm{i}, \mathrm{j} \in \mathbb{I}_N \equiv [-N/2, ..., N/2-1] \times [-N/2, ..., N/2-1]$ for some natural number N.

4 Pupil lattice

We are interested in the PSF that develops from a single incoming plane wave. We assume the intensity of the wave is uniform over the telescope entrance pupil, which leads us to use a particularly simple lattice to describe the entrance pupil plane, with primitive lattice vectors

$$\vec{u}_1 = \Delta u \, \hat{\imath} \tag{4}$$

and

$$\vec{u}_2 = \Delta u \, \hat{\jmath} \tag{5}$$

for some asserted sampling length Δu. (Note we are overloading the i and j characters again, this time with hats to indicate they are unit vectors.)

We wouldn't necessarily need a square lattice for the pupil – parallelograms would be okay, or even a hexagonal "lattice" (uniformity is important though) – but a square lattice is easy so we stick with that for now.

We interpret each point in the pupil lattice as an incoming ray. Because they represent a plane wave, the rays' directions of propagation are all identical, forming some angle with the optic axis. Furthermore, the rays' phases are identical along any plane perpendicular to their direction of propagation. The rays will propagate through the optical system of interest, each experiencing different degrees of phase delays and changes in propagation direction. If we can calculate or model these changes, then we can reinterpret the rays as plane waves near the focal plane and use Eq. 1 to determine the PSF at any location of interest.

5 Paraxial Optics

We'll first describe obtaining K using paraxial optics, which basically means we assume that the rays are close to the optic axis and $\tan\theta \approx \sin\theta \approx \theta$.

5.1 Focal Plane

For the Huygens construction, we are free to choose the points $\vec{x} \in X$ at which to evaluate the PSF, and, given a set of rays (= plane waves) can directly evaluate the sum Eq. 1 at these locations. Evaluating the sum directly can be slow, since it's linear in both the number of rays used and the number of evaluation points. In practice, we are usually interested in a square lattice of rays (with say N rays on a side) and square lattice of evaluation points (with M points on a side), so the direct sum is $O(M^2N^2)$. Fortunately, Eq. 1 can, with a few approximations, be manipulated into the form of a 2D Discrete Fourier Transform, which can be evaluated in $O(N^2 \log N)$ time, provided we are able to give up the freedom to choose both M and N independently.

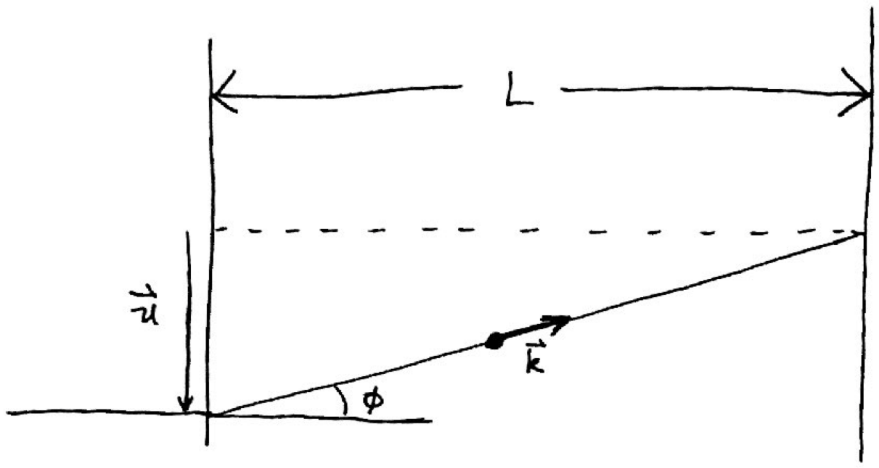

Figure 1: Relation between entrance pupil coordinate $\vec{u}$ and focal-plane-space wave vector $\vec{k}$ for a telescope with focal length L. A plane wave enters from the left where it encounters the entrance pupil, and gets focused to the right at the focal plane.

The approximation enabling the use of a Fourier transform concerns the relation between a ray's initial pupil coordinate $\vec{u}$ and its final wave vector component $\vec{k}$. Figure 1 illustrates this relationship under paraxial optics.

For pupil coordinates close to the optic axis (the paraxial approximation), we have

$$\frac{\vec{k}}{|\vec{k}|} = -\frac{\vec{u}}{\sqrt{|\vec{u}|^2 + L^2}} \approx -\frac{\vec{u}}{L} \tag{6}$$

or

$$\vec{k} \approx -\frac{2\pi}{\lambda n}\frac{\vec{u}}{L} \tag{7}$$

If we use this relation, we can construct a uniform rectilinear grid of $\vec{k}$ vectors (specifically, the x and y components of $\vec{k}$) that directly correspond the the initial rectilinear $\vec{u}$ vectors. I.e., the primitive lattice vectors for $\vec{k}$ are:

$$\vec{k}_1 = -\frac{2\pi}{\lambda n L}\vec{u}_1 \tag{8}$$

and

$$\vec{k}_2 = -\frac{2\pi}{\lambda n L}\vec{u}_2 \tag{9}$$

With this lattice, we can rewrite Eq. 1 as

$$A(\vec{x}) \propto \sum_{\mathrm{i,j}\in\mathbb{I}} P_{\mathrm{ij}} \exp\left(i\phi_{\mathrm{ij}}\right) \exp\left(i\vec{k}_{\mathrm{ij}} \cdot \vec{x}\right). \tag{10}$$

where P_{ij} indicates vignetted (0) or unvignetted (1) rays.

With the proper choice for a grid of $\vec{x}$ coordinates, we recognize the above as a 2D discrete Fourier transform.

For GalSim, the convention for the discrete Fourier transform between lattices with primitive vectors $\{\vec{a}_1, \vec{a}_2\}$ and $\{\vec{b}_1, \vec{b}_2\}$ (and N points each) are that

$$b_i a_i = \frac{2\pi}{N}, i \in \{1, 2\} \tag{11}$$

A slightly more general relation that allows for lattices with uniform but non-rectilinear grids is obtained from the reciprocal lattice formalism in solid-state physics. (E.g., see `https://physics.stackexchange.com/questions/340860/reciprocal-lattice-in-2d`) For GalSim DFTs, this becomes:

$$\vec{b}_i \cdot \vec{a}_j = \frac{2\pi}{N}\delta_{ij} \tag{12}$$

For paraxial optics though, we simply take focal plane space primitive lattice vectors resulting from Eq. 11:

$$\vec{x}_1 = -\frac{\lambda n L}{N\Delta u}\hat{\imath} \tag{13}$$

$$\vec{x}_2 = -\frac{\lambda n L}{N\Delta u}\hat{\jmath} \tag{14}$$

which allow us to write

$$A(\vec{x}_{\rm pq}) \propto \sum_{\rm i,j\in\mathbb{I}} P_{\rm ij} \exp\left(i\phi_{\rm ij}\right) \exp\left(i\vec{k}_{\rm ij} \cdot \vec{x}_{\rm pq}\right). \tag{15}$$

or

$$A(\vec{x}_{\rm pq}) \propto \left[\mathrm{DFT}\left(P_{\rm ij} \exp\left(i\phi_{\rm ij}\right)\right)\right]_{\rm pq} \tag{16}$$

5.2 Sky plane

The previous subsection tells us how to compute the PSF intensity under Fourier optics on the focal plane, where the dimensions on the focal plane are in length. We are also interested, however, in the PSF intensity distribution as projected on the sky. For this, we need the plate scale of the telescope.

Again, using a toy model, we can get an approximate relation between focal plane coordinates and sky coordinates. Figure 2 shows that this relation is

$$\vec{\theta} = \hat{\theta} \arctan\frac{|\vec{x}|}{L} \approx \frac{\vec{x}}{L} \tag{17}$$

where $\vec{\theta}$ is the angle on the sky (say, a 2d tangent plane projection), the telescope focal length is again L, and the focal plane coordinate is $\vec{x}$. We can thus use sky-plane primitive lattice vectors:

$$\vec{\theta_1} = -\frac{\lambda n}{N\Delta u}\hat{\imath} \tag{18}$$

$$\vec{\theta_2} = -\frac{\lambda n}{N\Delta u}\hat{\jmath} \tag{19}$$

The telescope focal length conveniently divides out.

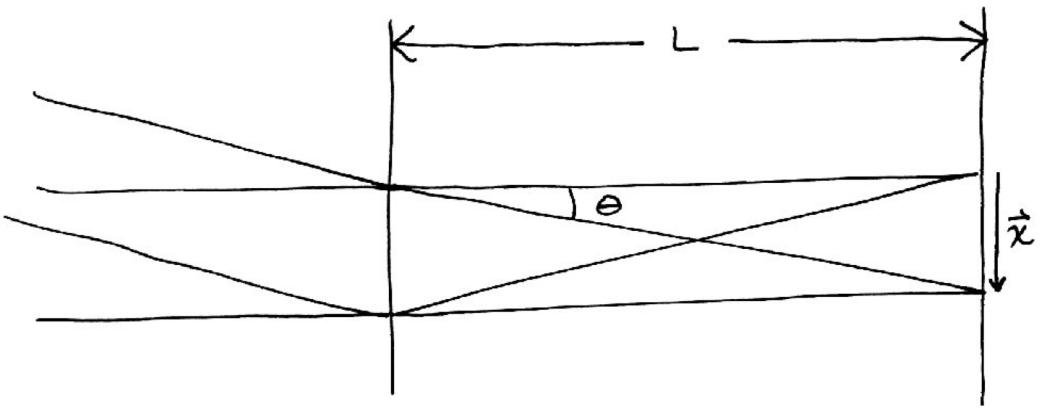

Figure 2: Toy telescope model for determining the plate scale.

Ignoring the refractive index n, this is equivalent to the GalSim convention of

$$\Delta\theta = \frac{\lambda}{N\Delta u} \tag{20}$$

We'll label the Fourier conjugate of the sky-coordinate $\vec{\theta}$ as $\vec{q}$. Using the same reciprocal lattice formalism as before, we find that the sky-plane Fourier lattice has primitive vectors

$$\vec{q_1} = -\frac{2\pi\Delta u}{\lambda n}\hat{\imath} \tag{21}$$

$$\vec{q_2} = -\frac{2\pi\Delta u}{\lambda n}\hat{\jmath} \tag{22}$$

Ignoring the refractive index, this is equivalent (up to a minus sign) to the GalSim convention of

$$\Delta q = \frac{2\pi}{\lambda}\Delta u \tag{23}$$

6 Non-paraxial optics

Paraxial optics are great fun, but don't tell the entire story, especially for a telescope with a wide field of view and a fast beam. Here, we look at more precise relationships between the various planes of interest, determined by tracing rays.

6.1 Focal plane

Leaving the grid on the pupil plane the same as in the paraxial case, our first difference for more precise optics concerns the relation between the pupil plane and the Fourier conjugate to the image plane; i.e., $\vec{k}(\vec{u})$.

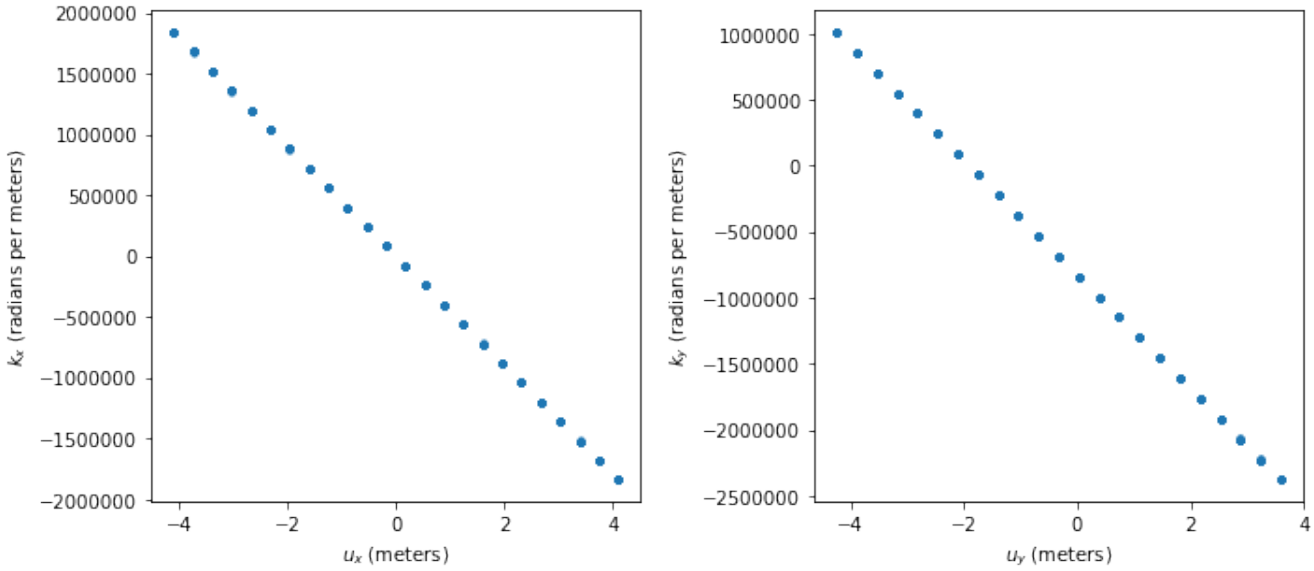

Figure 3: Relation between entrance pupil coordinate in meters and wave vector component at the focal plane in radians of phase per meter. The relationship is nearly linear. Note that the relationship does not pass through the origin in the y-direction, which is related to the degree of non-telecentricity of the Subaru/HSC system when the incoming plane wave is tilted in y.

Figure 3 shows the relationship between $\vec{k}$ and $\vec{u}$ for the Subaru/HSC telescope and camera determined using the raytracing software BATOID.

The incoming plane wave has wavelength $750nm$ and is tilted 0.01 radians to the optic axis in the +y direction.

To demonstrate just how linear the relations are, we show the residuals to bilinear fits of k_x and k_y against u_x and u_y in figure 4. The coefficients in the fit (ignoring the constant piece), are the Jacobian of the transformation between $\vec{u}$ and $\vec{k}$. For this particular situation, this Jacobian is:

$$\frac{\partial \vec{k}}{\partial \vec{u}} = \begin{bmatrix} \frac{\partial k_x}{\partial u_x} & \frac{\partial k_x}{\partial u_y} \\ \frac{\partial k_y}{\partial u_x} & \frac{\partial k_y}{\partial u_y} \end{bmatrix} = \begin{bmatrix} -4.49 \times 10^5 & 6.62 \times 10^{-2} \\ 2.11 \times 10^{-1} & -4.33 \times 10^5 \end{bmatrix} \tag{24}$$

(in units of radians of phase per meter squared).

The off-diagonal elements above ought to be zero for a perfectly symmetric optical system. The HSC obscuration pattern (from the spider) is not perfectly symmetric, however, so these terms show up as non-zero, though still subdominant to the diagonal entries.

The diagonal components are close to the predicted value from the paraxial approximation of

$$-\frac{2\pi}{\lambda n L} = -5.6 \cdot 10^5 \,\mathrm{rad\,m^{-2}} \tag{25}$$

where I used 15 m as the focal length (possibly too short). However, it's notable that the diagonal components are not the *same*. In particular, using $\vec{k} = \frac{\partial \vec{k}}{\partial \vec{u}} \vec{u}$ and Eq. 12, we find

$$\vec{x}_1 = \frac{1}{N\Delta u}(-1.40 \times 10^{-5} \hat{\imath} + 2.14 \times 10^{-12} \hat{\jmath}) \tag{26}$$

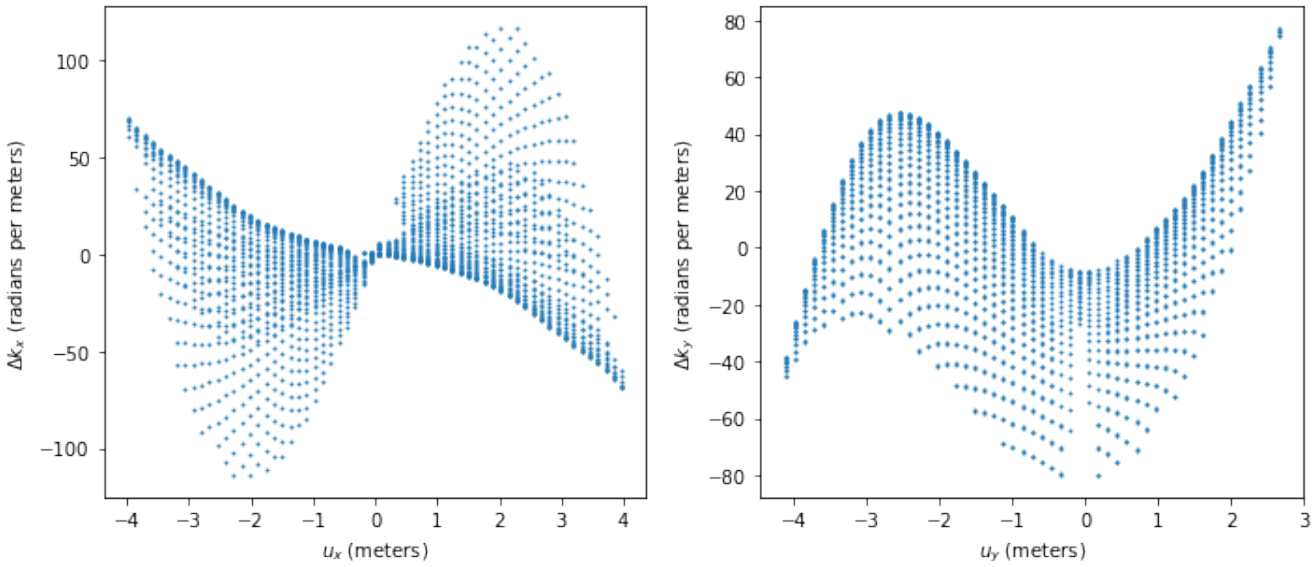

Figure 4: Relation between entrance pupil coordinate and the residual to a linear fit of the wavevector component against both pupil coordinate components. Note the y-axis scale is a factor of $\sim 10^4$ smaller than in the previous figure.

$$\vec{x}_2 = \frac{1}{N\Delta u}(-6.80 \times 10^{-12}\hat{\imath} - 1.45 \times 10^{-5}\hat{\jmath}) \tag{27}$$

6.2 Sky plane

To transform from focal plane coordinates to sky coordinates, we use BATOID to trace ray grids with slightly different incoming angles and look at where they focus on the focal plane to measure the Jacobian from a finite difference.

With this approach, and continuing the above example, I find gnomic primitives of

$$\vec{\theta}_1 = \frac{1}{N\Delta u}(-7.50 \times 10^{-7}\hat{\imath} + 1.13 \times 10^{-12}\hat{\jmath}) \tag{28}$$

$$\vec{\theta}_2 = \frac{1}{N\Delta u}(1.39 \times 10^{-11}\hat{\imath} + -7.50 \times 10^{-7}\hat{\jmath}) \tag{29}$$

(From the gnomic primitives, we ought to convert to real spherical coordinate primitives, but for the polar angles required here – no more than 1.75 degrees, even for LSST – the difference, aside from a rotation, is never more than about 0.1%.)

Notice that the diagonal elements are now equal, removing the distortion that appears in focal plane coordinates! In fact, these primitives are nearly indistinguishable from Eqs. 18 and 19. This result appears to be robust in my simulations, being reproducible for Subaru/HSC, LSST, and Blanco/DECam optics and for various field angles. I think this is likely a consequence of Liouville's theorem applied to optics (conservation of étendue), which roughly corresponds to

$$\delta\vec{u}\,\delta\vec{\theta} \propto \delta\vec{k}\,\delta\vec{x} \tag{30}$$

or equivalently

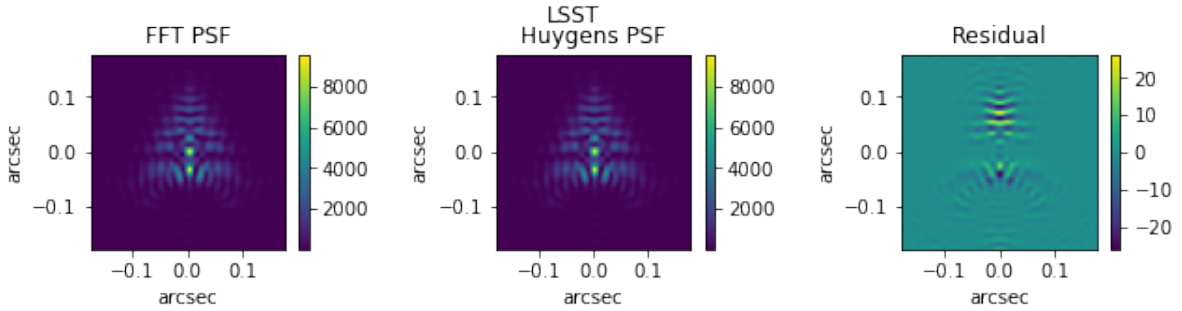

Figure 5: LSST PSF comparison for a point near the edge of the LSST field of view at 1.75 degrees off-axis. Left: PSF array determined using DFT algorithm. Center: PSF array determined from direct Huygens PSF sum. Right: FFT array minus Huygens array. The scales on the x- and y-axes are in arcseconds.

$$\frac{\partial \vec{k}}{\partial \vec{u}} \cdot \frac{\partial \vec{\theta}}{\partial \vec{x}} \propto \mathbb{1}_2 \tag{31}$$

The LHS of the above proportionality occurs in our transformation of primitive vectors from $U \rightarrow K \rightarrow X \rightarrow \Theta$.

At any rate, it appears that we can ignore complications of coordinate systems and WCS's if we simply interpret the optical PSFs generated by Fourier optics as living directly in sky coordinates, using Eqs. 18 and 19.

7 Fourier PSF vs. Huygens PSF

Our next task is to actually compare PSF surface brightness profiles generated by directly evaluating Eq. 1 and by using the DFT approximation Eq. 16 (This corresponds to alternately acknowledging or ignoring the small residuals shown in Figure 4). Since we do have freedom of where (in x) to evaluate the direct sum Huygens PSF, but not the FFT PSF, we will choose to evaluate the former at the points dictated by the latter to make the comparison.

For both algorithms, we evaluate the $\phi_{\vec{k}}$ by tracing rays to a reference sphere with center at the mean spot position on the focal plane and with radius approximately equal to the distance from the focal plane to the exit pupil. The $\phi_{\vec{k}}$ are the relative phases of the rays as they cross the reference sphere. We have checked that varying the reference sphere radius does not appreciably influence the relative phases $\phi_{\vec{k}}$.

Figures 5, 6, and 7 show comparisons between PSFs generated using the FFT algorithm and those generated using the direct summation algorithm. In all cases, the differences are on the order of a few parts per thousand of the peak surface brightness.

We're also interested in donuts. For this we simply defocus the telescope in the simulation (1.0 mm for LSST, 0.9 mm for HSC, and 1.5 mm for DECam) and repeat the above analysis. Figures 8, 9, and 10 show donut images generated with the FFT algorithm and the Huygens algorithm. The differences are slightly larger in this case: up to 5% of the peak surface brightness for LSST donuts, about 1% for HSC, and smaller for DECam. Note that positive residuals are often right next to negative residuals, especially for LSST, which means that if the donut model also included a convolution by a term to represent the atmosphere, they would get smoothed out somewhat.

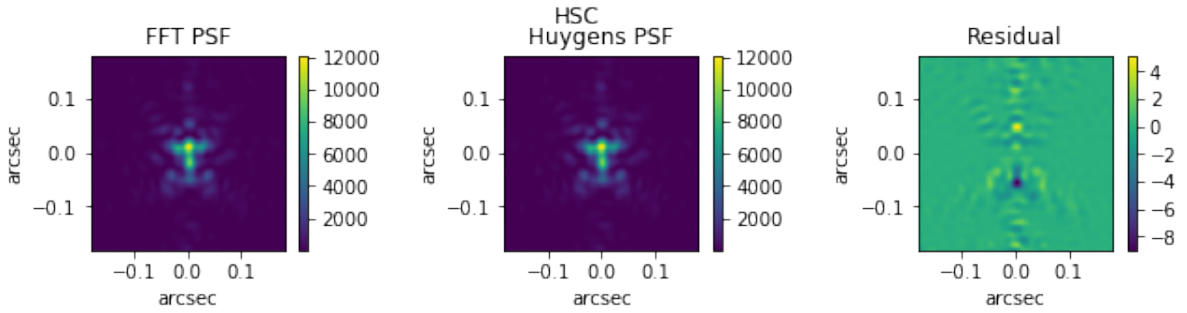

Figure 6: HSC PSF comparison for a point near the edge of the HSC field of view at 0.74 degrees off-axis.

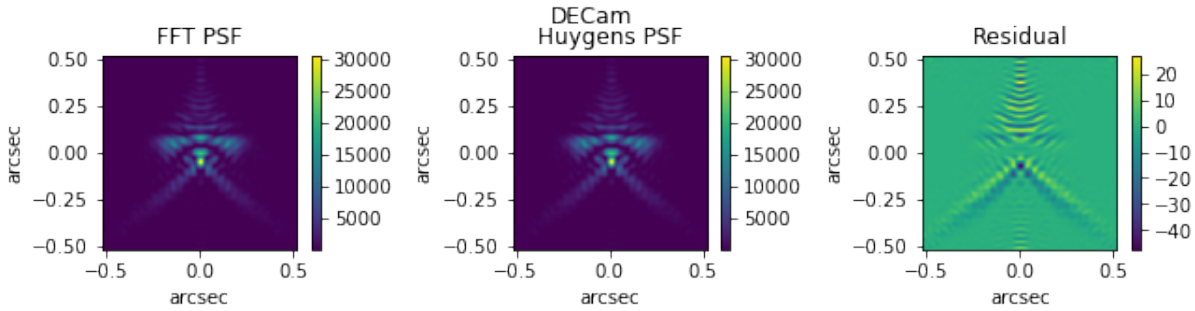

Figure 7: DECam PSF comparison for a point near the edge of the HSC field of view at 1.1 degrees off-axis.

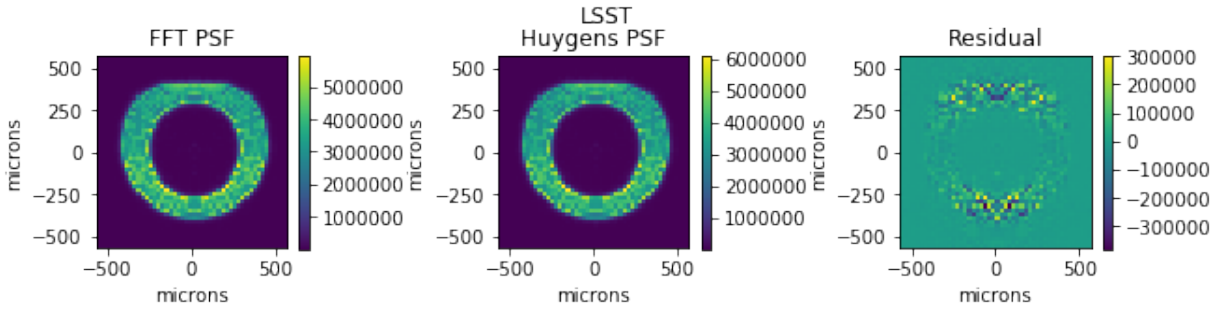

Figure 8: LSST donut comparison for a point near the edge of the LSST field of view at 1.75 degrees off-axis. Left: Donut determined using DFT algorithm. Center: Donut determined from direct Huygens PSF sum. Right: FFT donut minus Huygens donut. The scales on the x- and y-axes are in microns.

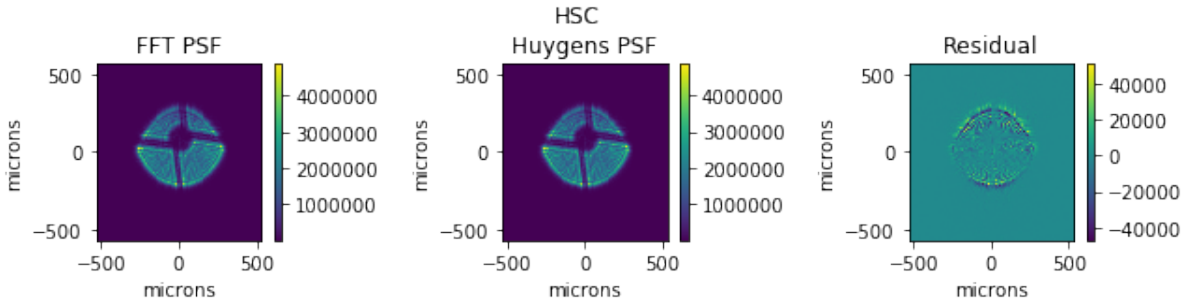

Figure 9: HSC donut comparison for a point near the edge of the HSC field of view at 0.74 degrees off-axis.

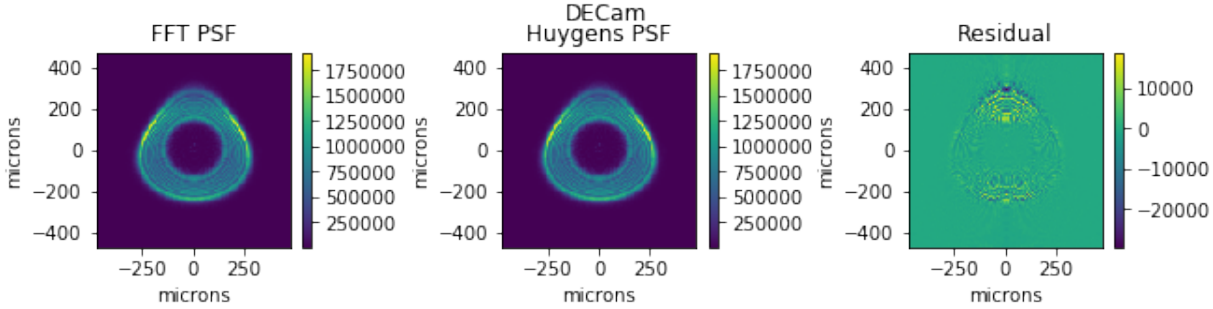

Figure 10: DECam donut comparison for a point near the edge of the HSC field of view at 1.1 degrees off-axis.

8 TLDR

Fourier optics generates PSF images in sky-coordinates following Eqs. 18 and 19. In focal plane coordinates, the transformations are more complicated, as they additionally incorporate distortion.

For LSST, HSC, and DECam, small discrepancies in the bilinear mapping from pupil coordinate to focal wave vector introduce up to few parts per thousand discrepancies in in-focus PSF surface brightness calculations, and up to few percent discrepancies in donut images. These discrepancies would likely be mitigated somewhat after further convolution with an atmospheric PSF.