A NEW SHEAR ESTIMATOR FOR WEAK-LENSING OBSERVATIONS

NICK KAISER
Institute for Astronomy, University of Hawaii, 2680 Woodlawn Drive, Honolulu, HI 96822; kaiser@hawaii.edu
Received 1999 April 2; accepted 2000 February 1

ABSTRACT

We present a new shear estimator for weak-lensing observations that properly accounts for the effects of a realistic point-spread function (PSF). Images of faint galaxies are subject to gravitational shearing, followed by smearing with the instrumental and/or atmospheric PSF. We construct a “finite-resolution shear operator,” which when applied to an observed image has the same effect as a gravitational shear applied prior to smearing. This operator allows one to calibrate essentially any shear estimator. We then specialize to the case of weighted second-moment shear estimators. We compute the shear polarizability, which gives the response of an individual galaxy’s polarization to a gravitational shear. We then compute the response of the population of galaxies, and thereby construct an optimal weighting scheme for combining shear estimates from galaxies of various shapes, luminosities, and sizes. We define a figure of merit (an inverse shear variance per unit solid angle) that characterizes the quality of image data for shear measurement. The new method is tested with simulated image data. We discuss the correction for anisotropy of the PSF and propose a new technique involving measuring shapes from images that have been convolved with a recircularizing PSF. We draw attention to a hitherto ignored noise-related bias and show how this can be analyzed and corrected for. The analysis here draws heavily on the properties of real PSF’s, and we include in Appendix A a brief review, highlighting those aspects that are relevant for weak lensing.

Subject headings: dark matter — galaxies: clusters: general — gravitational lensing — large-scale structure of universe

1. INTRODUCTION

In the weak-lensing or thin-lens approximation, the effect of gravitational lensing on the image of a distant object is a mapping of the surface brightness:

\[ f(r) = f(\delta_{ij} - \psi_j r_j) , \]

where the 2-vector \( r_j \) is the angular position on the sky measured relative to the center of the image, \( f \) is the intrinsic surface brightness that would be seen in the absence of lensing, and the symmetric “distortion tensor” \( \psi_j \) is an integral along the line of sight of the transverse second derivatives of the gravitational potential \( \Phi \) (Gunn 1967). In an open cosmology, for instance, the distortion for an object at the conformal distance \( \alpha_j \) can be written as

\[ \psi_{lm} = 2 \int \sin \omega \sinh \omega \sinh (\alpha_j - \omega) \sin \alpha \frac{\partial \alpha}{\partial x_i} \frac{\partial \alpha}{\partial x_j} \Phi \]

(Bar-Kana 1996; Kaiser 1998), where \( \partial_i \equiv \partial/\partial x_i \) and \( x_i = \theta_i \sin \omega \), where \( \theta_i \) is a two-component Cartesian vector in the plane of the sky and where the potential is related to the density contrast by \( V^2 \Phi = 4\pi G \delta \rho \), the Laplacian here being taken with respect to proper spatial coordinates. Equation (2) can be generalized to deal with sources at a range of distances, and with either accurately known redshifts or partial redshift information from broadband colors.

The distortion will, in general, change both the shapes and sizes, and hence the luminosities, of distant objects. Any component of the distortion that is coherent over large scales (larger than the typical angular separation of background galaxies) is therefore potentially observable as a relative modulation of the counts of objects or as a statistical anisotropy of the galaxy shapes, and this allows one to constrain the fluctuations in the total density, \( \delta \rho \), on the corresponding scales. Here we focus on an analysis of shape anisotropy, or “image shear,” although the methodology is readily extendable to include the effects of magnification.

While the effect of a shear on the sky surface brightness (eq. [1]) is rather simply stated, no completely satisfactory method for estimating the distortion has yet emerged. Perhaps ideally one would attack this problem using likelihood; that is, one would ask: what is the probability to observe a given set of background galaxy images given that they are drawn from some statistically isotropic unlensed parent distribution, as a function of the parameters \( \psi_{lm} \)? Unfortunately, this does not seem to be a particularly tractable problem. It is further complicated by the finite resolution of real observations, and by noise in the images. Instead, what has been done is to adopt some plausible shape statistic—typically some kind of central second moment—and then compute how this responds to a gravitational shear. We will now review the various different shear estimators that have been proposed, how they relate to one another, and what their limitations are.

1.1. Projection Matrices and the Shear Operator

As a preliminary, we introduce some mathematical formalisms that will simplify the analysis. Symmetric \( 2 \times 2 \) tensors, such as \( \psi_{lm} \), feature prominently in what follows. Such tensors have 3 real degrees of freedom. It is conventional to parameterize, for instance, the 3 real degrees of freedom of \( \psi_{lm} \) by the triplet comprising the convergence \( \kappa \) and the shear \( \gamma_\alpha \), \( \alpha = 1, 2 \), with

\[ \psi_{ij} = \begin{bmatrix} \kappa + \gamma_1 & \gamma_2 \\ \gamma_2 & \kappa - \gamma_1 \end{bmatrix} . \]

A simple way to convert between triplet and tensor components is to use the three constant \( 2 \times 2 \) “projection
The symmetrized products of pairs of these matrices, 
\[ [M_A M_B] = \begin{bmatrix} M_{00} & M_{10} & M_{20} \\ M_{10} & M_{00} & 0 \\ M_{20} & 0 & M_{00} \end{bmatrix} \]
where \( \delta_{AB} \) denotes the Kronecker delta symbol. It then follows that the useful identities from the contractions of products and triple products are
\[ M_{A} M_{B} = 2 \delta_{AB} , \]
\[ M_{A} M_{B} M_{C} = 2(\delta_{BC} \delta_{00} + \delta_{AC} \delta_{B0} + \delta_{AB} \delta_{C0}) . \]
Any symmetric tensor \( t_{lm} \) can be written as a linear combination of projection matrices with coefficients \( t_{ij} \), that is, \( t_{lm} = t_{ij} M_{ij}, \) and using equation (6) we have \( M_{Bim} t_{lm} = M_{Bim} M_{Alm} t_{ij} = 2 \delta_{AB} t_{ij}, \) so \( t_{ij} = \frac{1}{2} M_{Alm} t_{lm}. \) In this language, the convergence and the shear are the three components of the triplet representation of the distortion tensor: \( \kappa = \frac{1}{2} M_{Alm} \psi_{lm} = \psi_{0} \) and \( \gamma = \frac{1}{2} M_{Alm} \psi_{lm} = \psi_{0}. \) We adopt the convention that uppercase Roman indices range over \((0,1,2)\), while lowercase Greek symbols range over \((1,2)\), and that repeated indices are to be summed over. A two-component "polar" \( t_{ij} \) transforms under rotations as \( t_{ij} \rightarrow R_{ij}(2\theta) R_{ij}^{T}, \) while \( t_{ij} \) transforms as a scalar under rotations. An alternative and widely used formalism (Schneider, Ehlers, & Falco 1992) is to regard \( \gamma_{1}, \gamma_{2} \) as the real and imaginary parts of a complex shear, but we do not adopt that approach here.

It is also convenient to define a "shear operator," \( S_{ij} \), that generates the mapping of equation (1),
\[ f' = S_{ij} f . \]
At linear order in \( \gamma \), which will be valid for sufficiently weak shear, one can perform a first-order Taylor expansion of the right-hand side of equation (1), and \( S_{ij} \) becomes the differential operator
\[ S_{ij} = 1 - \gamma_{ij} M_{a} M_{ai} r_{i} \partial_{j} , \]
where \( \partial_{j} = \partial / \partial r_{j} \). This operator is rather similar to a rotation operator. An important question is the domain of validity of equation (9). The answer depends on the content of the image to which it is applied. For an image containing information only at spatial frequencies below some upper limit \( k_{\text{max}} \), this will be a good approximation provided that \( r \ll 1 / (\gamma k_{\text{max}}) \), so for finite shear, equation (9) will not apply for spatial frequencies \( \gtrsim 1 / (\gamma r) \). The limit here is the combination of frequency and distance from the origin such that a shear of strength \( \gamma \) corresponds to a local translation of about an inverse wavenumber. We would, however, expect equation (9) to correctly describe the effect of shear on the low-frequency behavior of an image, in the sense that if applied to an image that has had the high-frequency information removed, then the result will be essentially identical to the low-frequency content of an exactly sheared image.

1.2. Second-Moment Shear Estimators for Perfect Seeing

A pure shear will cause a circular object to become elliptical, and will change the ellipticities of noncircular objects. Following Valdes, Jarvis, & Tyson (1983), a natural choice of statistic to measure such a distortion is the second central moment, or quadrupole moment,
\[ q_{lm} = \int d^{2}r r_{i} r_{j} f(r) , \]
where we assume that the total flux, which is unaffected by a pure shear (at linear order), has been normalized such that \( \int d^{2}r f(r) = 1 \), and where the origin of coordinates has been taken such that the dipole moment \( d_{1} = \int d^{2}r r_{i} f(r) \) vanishes. The triple coefficients of the symmetric matrix \( q_{lm} \) are
\[ q_{A} = \frac{1}{2} M_{Aij} q_{ij} = \begin{pmatrix} (q_{xx} + q_{yy})/2 \\ (q_{xx} - q_{yy})/2 \\ q_{xy} \end{pmatrix} . \]
The first component, \( q_{00} \), is a measure of the size, or area, of the object, while the latter two, \( q_{x} \), are a measure of the eccentricity of the object (they vanish for a circular object), which we will refer to as the "polarization." We can compute how the quadrupole moment is affected by a shear by applying equation (9) to in equation (10) to find
\[ q_{lm} = q_{lm} - \gamma_{2} M_{aij} d^{2}r r_{i} r_{m} \partial_{j} f = q_{lm} + 2\gamma_{2} M_{aii} q_{lm} , \]
where we have integrated by parts and invoked the symmetry \( (M_{aii} = M_{aji}) \) and tracelessness \( (M_{a} = 0) \) of the matrices \( M_{1} \) and \( M_{2} \). With \( q_{lm} = M_{Bim} q_{B} \), etc., and using equation (7), we find
\[ \delta q_{A} = \frac{1}{2} M_{Alm}(q_{lm} - q_{lm}) = \gamma_{2} M_{Alm} M_{aii} M_{Bim} q_{B} = 2\gamma_{2}(\delta_{A0} q_{0} + \delta_{A} q_{0}) , \]
or equivalently,
\[ q'_{A} = q_{A} + 2\gamma_{2} q_{0} , \]
\[ q_{0} = q_{0} + 2\gamma_{2} q_{a} , \]
so the ratio of fractional changes in \( q_{A} \) and \( q_{0} \) is inversely proportional to the square of the ratio of \( q_{A} \) and \( q_{0} \). A weak gravitational shear therefore causes a change in the polarization, \( \delta q_{x} = 2\gamma_{2} q_{a} \), which is proportional to the area, \( q_{0} \), and it also induces a change in the area, \( \delta q_{0} \), which is proportional to the eccentricity. For intrinsically randomly oriented galaxies, the average intrinsic polarization, \( \langle q_{x} \rangle \), vanishes by symmetry, so \( \langle q_{x} \rangle = 2\langle q_{r} \rangle q_{a} \) and \( \langle q_{r} \rangle = \langle q_{0} \rangle \), and therefore
\[ q_{a} = \frac{\langle q_{r} \rangle}{2\langle q_{0} \rangle} , \]
which one can use to estimate the shear on a patch of sky by replacing the averaging operator \( \langle \ldots \rangle \) by a summation, \( \sum \ldots / N \). This type of shear estimator was first introduced and used by Valdes et al. (1983) in their search for large-scale shear. The averages of \( q_{A} \) and \( q_{0} \) here are heavily weighted toward larger galaxies, which is not ideal. More commonly, what has been done (e.g., Tyson et al. 1984) is to normalize the polarization by the trace of the second moments and
define an “ellipticity vector”
\[ e_s = \frac{q_x}{q_0}, \] (16)
which depends only on the galaxy shape, and whose expectation value is
\[ \langle e_s^2 \rangle = \frac{q_x}{q_0 + 2q_0} \approx 2\gamma_q - 2\gamma_q \left( \frac{q_x}{q_0} \right) \]
\[ = 2\gamma_q \left( 1 - \gamma_q \right), \] (17)
and where we have kept only terms up to linear order in the shear (Kaiser, Squires, & Broadhurst 1995, hereafter KSB).

An alternative (Bonnet & Mellier 1995) is to normalize the second moments by the square root of the determinant of \( q_{lm} \) rather than by the trace:
\[ e_{s,\text{BM}} = \frac{q_x}{\sqrt{|q_{lm}|}} = \frac{q_x}{\sqrt{q_0^2 - q_x q_y}}, \] (18)
which, again at first order in \( \gamma \), has the expectation value
\[ \langle e_{s,\text{BM}}^2 \rangle = 2\gamma_q \left( 1 + \gamma_q \right) e_{s,\text{BM}}^2. \] (19)

Yet another possibility is to use only the information in the position angle, \( \theta = (\tan^{-1} q_2/q_1)/2 \) (Kochanek 1990), or, equivalently, in the unit shear vector, \( |e_s| \); there are numerous other similar statistics one could use.

These formulae for the response of the polarization statistics should be used with caution. While formally correct, the averages here must be taken in an unweighted manner over all galaxies. This is generally neither possible nor particularly desirable, since there are detection limits, and one would ideally like to estimate the shear as some optimally weighted combination of polarizations of galaxies of different fluxes and sizes, and these simple relations no longer hold. We will return to this later. First, however, let us consider the effect on these estimators of a finite point-spread function, which is well illustrated by the simple estimator of equation (15).

1.3. Second-Moment Shear Estimators for Finite Seeing

In real observations, some or all of atmospheric turbulence, optical aberrations, aperture size, guiding or registration errors, atmospheric dispersion, finite pixel size, scattering, etc., will combine to give an observed surface brightness of
\[ f_o(r) = \int d^2r' g(r') f(r - r') \equiv g \otimes f, \] (20)
where \( g(r) \) is the point-spread function (PSF). The combined effect of gravitational shearing, followed by instrumental and atmospheric seeing, is the transformation
\[ f_o' = g \otimes S_r f. \] (21)

Circular seeing will tend to reduce the ellipticity, while departures from circularity will introduce an artificial polarization. To make accurate shear measurements, we need to correct for the latter and calibrate the former. For estimators such as equation (15), which are computed from moments \( q_{lm} \) as defined in equation (10), the effect of the PSF is rather simple, since, as noted by Valdes et al. (1983), the second central moment of a convolution of two normalized functions is just the sum of the second central moments,
\[ q_{lm}(f_o) = q_{lm}(f) + q_{lm}(g), \] (22)
and this additive property is shared by the independent components \( q_4 \). Thus, one can recover the second moments, \( q_{lm}(f) \), that would be measured by a large perfect telescope in space from the observed moments, \( q_{lm}(f_o) \), simply by subtracting the moments of the PSF, \( q_{lm}(f) = q_{lm}(f_o) - q_{lm}(g) \). These can be measured from shapes of foreground stars in the image, or, in the case of diffraction-limited seeing, computed from the telescope design. In terms of observed moments, the shear estimator equation (15) becomes
\[ \gamma_s = \frac{q_x - q_y}{2(q_0 - q_0^*)}, \] (23)
where the asterisk denotes the value for a stellar object. This procedure—using equation (22) to correct the measured moments of the PSF—simply and rather elegantly compensates for both the circularizing and the distorting effects of realistic seeing.

1.4. Weighted-Moment Shear Estimators

Unfortunately, while very simple to analyze, the shear estimator constructed from the second moments as defined in equation (10) is not at all useful when applied to real data. For one thing, photon-counting noise introduces an uncertainty in the moments that diverges as the square of the radius out to which one integrates, and the effect of neighboring objects will similarly grossly corrupt the signal. There is also the problem that for the kinds of PSFs that arise in real telescopes, the second moment is not well defined. To obtain a practical estimator, it is necessary to truncate the integral in equation (10). This can be done in various ways; the approach implemented in the FOCAS software package (Jarvis & Tyson 1981) and also in the SEXTRACTOR package (Bertin & Arnouts 1996) is to truncate the integral at some isophotal threshold \( f_{crit} \) and compute moments of the nonlinear function \( f_s(r) = \Theta(f_{crit} - f_{run})f_r \), where \( \Theta(f) \) is the Heaviside function. In fact, isophotal moments are most commonly computed from an image that has been smoothed with some kernel (usually an approximate model of the PSF itself), so that the isophotal boundary will be well defined. An alternative, as advocated by Bonnet & Mellier (1995) and KSB, is to limit the range of integration with a user-supplied weight function, \( w(r) \), and define
\[ q_{lm} \propto \int d^2r w(r)r_rr_m f_s(r). \] (24)

Another possibility is to define a polarization vector in terms of the second derivatives of a smoothed image, \( f_s = w \otimes f \), as
\[ q_{lm} \propto \delta_{lm} f_s, \] (25)
evaluated at the peak of \( f_s \). For a Gaussian weight function \( w(r) \), however, this is essentially equivalent to the weighted second moment (eq. [24]). As we shall see, the weighted-moment statistics, being linear in the surface brightness, offer significant advantages over the isophotal threshold.
method, but in either case, the simple relation given in equation (22) between observed and intrinsic second moments no longer holds, and compensating for the effects of the PSF becomes considerably more complicated than equation (23).

A partial solution to this problem has been offered by KSB, who computed the response of shear estimators such as equation (16) to an anisotropy of the PSF under the assumption that this can be modeled as the convolution of a circular PSF, \( g_{\text{circ}}(r) \), with some compact but possibly highly anisotropic function \( k(r) \). This would be a reasonable approximation, for example, for the case of atmospheric seeing in the presence of small-amplitude guiding errors. They found that such a PSF anisotropy would introduce an artificial ellipticity, \( \delta \epsilon = P_{\text{sm}}(f_{\text{a}}) \psi_p \), where \( \psi_p = M_{\text{gal}} \int d^2r r r_m k(r) \) is the unweighted polarization of \( k(r) \) (although see Hoekstra et al. 1998, who corrected a minor error in the analysis). The “smear polarizability,” \( P_{\text{sm}} \), is a combination of weighted moments of \( f_{\text{a}} \), and is essentially a measure of the inverse area of the object. An interesting feature of this analysis is that only the second moment of the convolving kernel appears here; all other details of \( k(r) \) are irrelevant, and it is relatively straightforward to determine \( p_f \) from observed stars, set up a model for how this two-vector field \( p_f(r) \) varies across the field, and then correct the ellipticities, at least at linear order, to what would have been measured by a telescope with a perfectly circular PSF.

KSB also computed the response of the ellipticity to a shear applied to \( f_{\text{a}} \) after smearing with the PSF, and found

\[
\delta \epsilon_c = P_{\text{sm}}(f_{\text{a}}) \gamma_p q_g, \quad \text{where } P_{\text{sm}} \text{ is another combination of moments of } f_{\text{a}}.
\]

This of course, is not what one wants, since one really needs the response to a shear applied before smearing with the PSF, and KSB suggested that one use deep images from the Hubble Space Telescope (HST) to empirically deduce a correction for finite seeing. A related approach, suggested by Wilson, Cole, & Frenk (1996), is to iteratively deconvolve the images, apply a shear, and then reconvolve and again use the change in the polarization with applied shear to calibrate the relation between \( \gamma \) and the polarization measured from the original images.

Luppino & Kaiser (1997) have used a somewhat different approach. They note that the real operation equation (21) can be written as \( g \otimes S, f = S,(S^{-1} g) \otimes f \), i.e., applying a shear before smearing is equivalent to smearing with an antisheared PSF \( S^{-1} g \) and then shearing. If the PSF is Gaussian, applying a weak shear to it is precisely equivalent to smearing it with another small but anisotropic Gaussian; the effect of this can therefore be computed using the smear polarizability of KSB, and it then follows that for a nearly circular Gaussian PSF, equation (21) will cause a response

\[
e_\gamma \to e' = e_\gamma + \delta e_\gamma = e_\gamma + P_{\text{sm}}(f_o) \gamma_p q_g + P_{\text{sm}}(S^{-1} g), \quad (26)
\]

where \( p_{\text{sm}}(S^{-1} g) \) is the unweighted polarization of the anti-sheared PSF and is of first order in \( \gamma \). One can most easily infer the value of \( p_{\text{sm}}(S^{-1} g) \) from the values of the shear and smear polarizabilities for a stellar object: shearing a point source has no effect, so for a star we must have \( p_{\text{sm}} = -[P(\gamma)/P_{\text{sm}}(\gamma)] \gamma_p q_g \), where we have suppressed the indices on the stellar polarizabilities (for a nearly circular PSF these are approximately diagonal), and hence the net effect of a real shear in this approximation is

\[
\delta e_\gamma = \left[ P_{\text{sm}}(f_o) - \frac{P(\gamma)}{P_{\text{sm}}(\gamma)} P_{\text{sm}}(f_o) \right] \gamma_p, \quad (27)
\]

This is nice, since it expressed the linear response of the polarization, \( e_\gamma \), to a shear entirely in terms of the observables \( f_o \) and \( g \), but it rests somewhat shakily on the assumption that shearing a real PSF can indeed be modeled as shearing it with some compact kernel. For a Gaussian PSF this is exact, but that is a rather special, and unfortunately unrealistic, case.

One indication that equation (27) cannot apply for a general PSF comes from considering the factor \( P(\gamma)/P_{\text{sm}}(\gamma) \). At no point have we specified the scale of the weighting function \( w(r) \), so this factor must be invariant to choice of scale length. For a Gaussian PSF this is indeed the case, but for a PSF generated by atmospheric turbulence, for instance, this is not the case, and equation (27) is then inconsistent. Similarly, the factor \( q_g \) appearing in the KSB correction for PSF anisotropy is, in general, dependent on the scale of the window used to measure the PSF properties. Hoekstra et al. (1998) have found from an analysis of simulated images that for the HST WFPC2 instrument, one can adjust the scale of the weight function for the stars to render the calibration equation (27) and the KSB anisotropy correction reasonably accurate, but it is not clear that this will apply in general. Indeed, for diffraction-limited seeing, the inadequacy of the KSB formalism has a deeper root. While the assumption that the real PSF can be modeled as a convolution of a circular PSF with some compact kernel \( k(r) \) may be a good approximation for atmospheric-turbulence seeing with small-amplitude guiding errors and such, as we shall see, this is not the case in general.

The current situation is therefore somewhat unsatisfactory. In this paper we develop an improved method of shear estimation that does not suffer from the inadequacies noted above and works for a PSF of essentially arbitrary form. The layout of the paper is as follows. In § 2 we construct an operator that generalizes equation (9) to finite-size PSFs and generates the effect of a gravitational shear on the observed (i.e., postseeing) surface brightness, \( f_o \). We first show that, quite generally, the effect on \( f_o = g \otimes f \) of a shear applied to \( f \) is equivalent to a shear applied directly to \( f_o \) plus a “commutator” term, which is a convolution of \( f_o \) with a kernel \( S(\gamma_H f) \), where \( H(\gamma) \) can be computed from the PSF \( g \). We explore the properties of this kernel for various types and PSFs. The kernel \( H(\gamma) \) involves \( \gamma_H \) (where the tilde denotes the Fourier transform), which appears, particularly in the case of diffraction-limited seeing, to be formally ill defined, since \( \gamma_H \) vanishes at finite radius. We show, however, that the operator generating the effect of a shear on a filtered image that has had frequencies close to the diffraction limit attenuated does not suffer from this problem. In § 3, we specialize to weighted second moments and compute their response to shear, both for individual objects (§ 3.1) and for the population of galaxies of a given flux, size, and shape (§ 3.2). In § 3.3 we show how to optimally combine estimates of the shear from galaxies of various different types. The method is tested using simulated mock images. In § 4 we discuss the correction for PSF anisotropy, and propose a new technique involving measuring shapes from images that have been convolved with a recircularizing PSF. We draw attention to a hitherto ignored noise-related bias and show how this can be analyzed and corrected for. In § 5 we summarize the main results and outline how they can be applied to real data. In the analysis here, we draw heavily on the properties of real
physical PSFs, and we include in Appendix A a brief review of basic PSF theory and discussion of PSF properties as they relate to weak-lensing observations.

2. FINITE-RESOLUTION SHEAR OPERATOR

We now consider how the shear operator given in equation (9) is modified by finite resolution arising in either the atmosphere, the telescope optics, or the detector. Let the unperturbed surface brightness (i.e., that which would have been observed in the absence of lensing) be

$$f_{\theta} = g \otimes f,$$

so the perturbed surface brightness is

$$f'_{\theta} = g \otimes S_{\gamma} f.$$  

Fourier transforming and using the result that, at linear order in $\gamma$, applying a shear in real space is equivalent to applying a shear of the opposite sign in Fourier space [i.e., if $a = S \cdot b$ then $\tilde{a} = S_{-\gamma} \tilde{b}$, where $\tilde{F}(k) = \int d^2r \, F(r)e^{i kr}$], we have

$$\tilde{f}'_{\theta} = \tilde{g} \otimes \tilde{S}_{\gamma} \tilde{f} - \tilde{g} \delta S_{\gamma} \left( \frac{\tilde{f}}{\tilde{g}} \right),$$

where $\delta S_{\gamma} \equiv S_{-\gamma} - 1$. Since $\delta S_{\gamma}$ is a first-order differential operator, we have $\delta S_{\gamma} \tilde{g} = \tilde{g}^{1-1} \delta S_{\gamma} \tilde{f} - \tilde{g}^{1-1} \delta S_{\gamma} \tilde{g}$, and $\tilde{g}^{1-1} \delta S_{\gamma} \tilde{g} = \delta S_{\gamma} \ln \tilde{g}$. Consequently, to first order in $\gamma$,

$$\tilde{f}'_{\theta} = \tilde{f} - \delta S_{\gamma} \tilde{f}_{\theta} + \delta S_{\gamma} \ln \tilde{g},$$

or, in real space,

$$f'_{\theta} = f_{\theta} + \delta S_{\gamma} f_{\theta} - (\delta S_{\gamma} h) \otimes f_{\theta},$$

where $h$ is the inverse transform of $\tilde{h} \equiv \ln \tilde{g}$, i.e., the logarithm of the optical transfer function (OTF). Invoking the definition of the shear operator given in equation (9), we have

$$f'_{\theta} = f_{\theta} - \gamma_{\theta} M_{\theta\bar{\theta}}[r_{\theta} \partial_{\bar{\theta}} f_{\theta} - (r_{\bar{\theta}} \partial_{\theta} h) \otimes f_{\theta}].$$

Note that the PSF is a real function, so $\tilde{g}(-k) = \tilde{g}^*(k)$, and this symmetry is shared by $\ln \tilde{g}$, so $h$ is also real.

The finite-resolution shear operator given by equation (33) is then

$$f'_{\theta} = S_{\gamma} f_{\theta} - (\delta S_{\gamma} h) \otimes f_{\theta},$$

that is, it is the regular shear operator, $S_{\gamma}$, applied to the postseeing image, $f_{\theta}$, plus a “commutator term,” $g \otimes S_{\gamma} f - S_{\gamma}(g \otimes f)$, which is a correction for finite PSF size and which is a convolution of $f_{\theta}$ with a kernel $\gamma_{\theta} H_{\gamma}(r) = \gamma_{\theta} M_{\theta\bar{\theta}}[r_{\theta} \partial_{\bar{\theta}} h]h = \delta S_{\gamma} h$, which one can compute from the PSF $g$. This seems quite promising; the effect of the first term on the polarization of an object is just $\gamma_{\theta}$ times the KSB postseeing shear polarizability. The response of the polarization to the second term should also be calculable; if the kernel is very compact, then the response will be given by the KSB linearized smear polarizability, but even if it is not, equation (33) should still allow one to compute the polarization response, since it expresses the change in $f_{\theta}$, and therefore in the polarization, or indeed any other statistic computable from $f_{\theta}$ directly in terms of the observed surface brightness, $f_{\theta}$ itself. The finite-resolution shear operator involves the function $h(r)$, which is the transform of the logarithm of the OTF. Since the OTF becomes exponentially small or may vanish, two questions immediately arise: Is the function $h(r)$ mathematically well defined? and can it be reliably computed from PSFs measured from real stellar images? To address these questions, we now explore the form of the commutator kernel, $H_{\theta}(r)$, for various types of PSF.

2.1. Gaussian Ellipsoid PSF

Consider first the simple (although unrealistic) case of a Gaussian ellipsoid PSF: $g(r) = (2\pi)^{-1/2} \exp(-r_{\theta} m_{\theta\theta}^{-1} r_{\theta}/2)$, where $m_{\theta\theta} = \langle r_{\theta} r_{\theta} \rangle$ is the matrix of second moments. In this case, $H_{\theta}(r) = -k_{\theta} m_{\theta\theta}^{-1} r_{\theta}/2$, and so $S_{\theta}$ in $\tilde{g}(k) = g_{r} M_{\theta\theta} m_{\theta\theta}^{-1} k_{\theta} k_{\theta}$; on transforming this we find that the kernel is the operator $H_{\theta}(r) = \gamma_{\theta} m_{\theta\theta}^{-1} \partial r_{\theta} \partial_{\theta}$ [in which case the function $H_{\theta}(r)$ can be realized, for example, as the contraction of the constant matrix $\gamma_{\theta} M_{\theta\theta}$ with the limit as $\sigma \to 0$ of the matrix of second partial derivatives of a Gaussian ball $(2\pi \sigma^2)^{-1} \exp(-r^2/2\sigma^2)$], and therefore

$$f'_{\theta} = f_{\theta} - \gamma_{\theta} M_{\theta\theta}[r_{\theta} \partial_{\bar{\theta}} f_{\theta} + m_{\theta\theta}^{-1} \partial r_{\theta} \partial_{\theta} f_{\theta}].$$

Therefore, for a Gaussian PSF, the finite-resolution shear operator is well defined and is purely local differential operator. Gaussian PSF’s are, however, unphysical, and do not arise in real instruments.

2.2. Atmospheric Turbulence

Now consider atmospheric turbulence–limited seeing. As reviewed in Appendix A, in that case $\tilde{g}(k) = \exp[-S(k\lambda 2\pi)/2]$, where for fully developed Kolmogorov turbulence, the structure function is $S(r) = 6.88(r_0/k)^{5/3}$, where $r_0$ is the Fried length, so in this case $\tilde{g} \propto \ln \tilde{g} \propto -r^{5/3}$. The commutator kernel therefore involves the transform of this power law, which diverges strongly at high $k$; how do we make sense of this? Dimensional analysis would suggest that $h(r) = \int d^2 k \, k^{5/3} e^{ikr}$ is a power law with $h(r) \propto r^{-11/3}$. The same argument, however, applied to a Gaussian would suggest $h(r) \propto r^{-4}$, which we know to be false, since we have just shown that in that case, $h(r)$ is just the second derivative of a $\delta$-function and has no extended tail. To clarify the situation, and to verify the validity of the power-law form for $h(r)$ for atmospheric seeing, consider the function $h(r; R)$, which is the transform of $k^{5/3}$ times an exponential cutoff function $\exp(-k R)$, that is,

$$h(r; R) \propto \int d^2 k \, k^{5/3} e^{-k R},$$

so $h(r)$ is the limit as $R \to 0$ of $h(r; R)$. If we write this as an integral with respect to the rescaled wavenumber $y = k R$, it becomes clear that $h(r; R)$ obeys a self-similar scaling with respect to choice of $R$ and can be expressed in terms of some universal function $F(y)$ such that $h(r; R) = R^{-11/3} F(r)$. If we postulate that $h(r; R)$ tends to some $R$-independent limit for finite $r$ as $R \to 0$, then that limit must be a power law with $F(z) \propto z^{-11/3}$, so $h(r)$ must be proportional to $r^{-11/3}$. This argument does not indicate the coefficient multiplying the power law, which happens to vanish for a Gaussian. The value of $h$ at the origin is $h(0; R) = 2 R^{-11/3} \Gamma(8/3)$, i.e., on the order of the value of the $r \gg R$ power-law asymptote extrapolated to $r \sim R$ (note that for the Gaussian, the analogous gamma function is not defined). A numerical integration for various values of $R$ is shown in Figure 1. This shows that as we decrease $R$, the function $h(r; R)$ does indeed tend to a $r^{-11/3}$ power law with finite, nonzero, $R$-independent amplitude, but that the power law breaks (with a change of sign) at $r \sim R$ and becomes asymptotically flat for $r \ll R$. For finite $R$, this function is well characterized as a positive
“softened δ-function” core with width \( \sim R \) and central value \( \sim R^{-11/3} \), and therefore with weight proportional to \( R^{-5/3} \), surrounded by a negative power-law halo with \( h \propto r^{-11/3} \). In the limit \( R \to 0 \), the core shrinks and becomes negligible, leaving only the power-law halo. This power law has the same slope as the well-known large-angle \( g(r) \propto r^{-11/3} \) form of the PSF, but the PSF departs from this law at the angular scale \( r_a \sim \lambda/2\theta_0 \) corresponding to the Fried length, whereas \( h(r) \) is a perfect power law and shows no features at the Fried scale.

The function \( h(r) \) diverges strongly at the origin, and the same is true of the kernel \( H_a = M_{a ij} \partial_i \partial_j h \), so \( H_a = -11/3r^{-11/3} \cos 2\varphi, \sin 2\varphi \). This does not give rise to any inconsistency with equation (33), however. The observed sky \( f_o \) is coherent on the scale of the PSF \( r_g \), so in computing the contribution to the commutator term

\[
\delta_o = \gamma_a M_{a ij} \int d^2 r' r_i' r_j' (r')^{-17/3} f_o (r - r')
\]

at small \( r' \ll r_0 \), we can perform a Taylor expansion of \( f_o \), and we find that the first nonvanishing term is the second-order term \( \delta_o \sim \langle \partial_i \partial_j f_o \rangle \int d^2 r' \left( \partial_i r_i' \right) \left( \partial_j r_j' \right) (r')^{-11/3} \), which has no physical divergence at \( r' = 0 \). Somewhat unfortunately, perhaps, the same line of argument shows that the commutator term cannot in this case be assumed to be a convolution with some compact function \( k(r) \). A necessary condition for this to be valid is that the unweighted second moment of the kernel \( p_a = M_{a ij} \partial_i \partial_j H_a \) should be well defined and tend to a finite limit within some small radius \( \ll r_g \), the characteristic width of the PSF. But this is the same integral as above, which does not tend to any well-defined value, but rather diverges as the 1/3 power of the upper limit on the integration radius. Thus, while the \( h(r) \propto r^{-11/3} \) asymptote is quite steep, it is not sufficiently steep to render the \( p_a \) value well defined. This strictly invalidates the approximation of Luppino & Kaiser (1997) for turbulence-limited seeing.

The Kolmogorov law is only expected to apply over a finite range of scales. The structure function will fall below the \( r^{5/3} \) form at the “outer scale,” which recent measurements at La Silla suggest to be typically on the order of 20 m (Martin et al. 1998; although see also the review of earlier results in Avila et al. 1997). Fast guiding (often referred to as “tip-tilt” correction) would also effectively reduce the structure function at these frequencies, and such effects act in a manner similar to the simple exponential cutoff we have assumed. To compute the OTF properly, one should include the effect of the aperture. The detailed form of \( h(r) \) at very small \( r \) will be sensitive to the detailed form of the outer scale cutoff and/or aperture, but provided that these lie at scales much greater than the Fried length (which is the case for large-aperture telescopes at good sites), the effect on the commutator term should be almost independent of the cutoff, because any signal at the relevant spatial frequencies will have been attenuated by a large factor. There will also be deviations from the Kolmogorov structure function law at small separation; diffusion will damp out fine-scale turbulence, and mirror roughness will add additional high spatial frequency phase errors. These effects will modify the \( h \propto r^{-11/3} \) halo at large \( r \), just as they modify the large-angle \( g \propto r^{-11/3} \) form of the PSF. These effects may profoundly influence the behavior of \( h, H_a \) at large angle, but have little impact on the type of shape statistics considered here, where the large-angle contribution to the polarization is suppressed by the weight function or the isophotal cutoff.

2.3. Diffraction-Limited Seeing

Let us now consider diffraction-limited observations. In this case, the OTF \( \tilde{g} \) falls to zero at finite spatial frequency (the diffraction limit) and so \( \ln \tilde{g} \) diverges as one approaches the diffraction limit from below and is not defined for higher frequencies. It is easy to understand how this divergence arises physically. As noted above, applying a weak shear in real space is equivalent to applying a shear of the opposite sign in Fourier space. Thus, the information in some Fourier mode of the sheared image comes from a slightly displaced mode in the unsheared image. If the OTF \( \tilde{g} \) is finite and continuous, as is the case for atmospheric turbulence–limited seeing, then the image \( f_o \) contains all of the information required to predict \( f_o' \). For diffraction-limited seeing, however, and for observations through a narrowband filter, the OTF has a well-defined edge, so the information contained in \( f_o' \) for spatial frequencies just inside the cutoff may lie outside the cutoff in \( f_o \), and so will be missing. For a finite bandpass, the form of the OTF will be modified, but must still fall to zero at the diffraction limit for the highest frequencies passed by the filter, and \( \ln \tilde{g} \) is still formally divergent.

Thus, in general, from knowledge of \( f_o' = g \otimes f \) it is strictly speaking not possible to say how \( f_o' \) would change in response to a small but finite shear applied before seeing; there are Fourier modes within a distance \( \delta k \sim k_{max} \) of the diffraction limit that one cannot predict. As an extreme example, imagine that we have a pure sinusoidal ripple on the sky that lies just outside the diffraction limit and that is therefore invisible. Applying an appropriate shear can bring that mode inside the limit and the ripple will appear as if from nowhere. As applied to real signals, however, this formally divergent behavior does not present a serious problem. First, if we observe galaxies of overall extent \( r_g \), which is typically not much larger than the seeing disk, then the transform must vary smoothly with the coherence scale, \( \delta k \sim 1/r_g \), which for sufficiently small \( \gamma \) will greatly exceed \( k_{max} \); this rules out the possibility of isolated spikes lurking just beyond the diffraction limit as in the example. In addition, for filled-aperture optical telescopes the marginal
modes are quite strongly attenuated, and for any finite measurement error from, e.g., photon-counting statistics, will contain very little information. The situation here is similar to that encountered in analyzing atmospheric turbulence PSF; there, while the very small scale details of $h(r)$ are sensitive to the aperture or outer scale cutoff, when applied to real data they have essentially no effect. Similarly, if one can generate a function that accurately coincides with $h = \log \tilde{g}$ where $\tilde{g}(k)$ exceeds some small value, but that tapers off smoothly at larger frequencies (rather than diverging at $k_{\text{max}}$), then this should have a well-defined transform and should give what is, for all practical purposes, a good approximation to the true finite-resolution shear operator.

One way to explicitly remove the divergence is to compute shapes from an image that has had the marginally detectable modes attenuated. If one convolves the observed field $f_0$ with some filter function $g^\prime(r)$ to make a smoothed image, $f_s = g^\prime \otimes f_o$, then in Fourier space we have

$$\tilde{f}_s = \tilde{f}_0 - \tilde{g}' \delta S \tilde{f}_0 + \tilde{f}_0 \tilde{g}' \delta S \tilde{g}, \quad (37)$$

so provided that $\tilde{g}'$ falls off at least as fast as $\tilde{g}$ as one approaches the diffraction limit, this operator is well defined. In real space, the corresponding operator is

$$f'_s = f_s - \gamma_s M_{sl} [g^\prime \otimes (r_1 \partial_{r_1} f_o) - g^\prime \otimes (r_1 \partial_{r_1} h) \otimes f_o], \quad (38)$$

so now the commutator term is the convolution of $f_s$ with $H_s = M_{sl} g' \otimes (r_1 \partial_{r_1} h)$. In principle, one can design $g'$ such that $f_s$ is essentially identical to $f_o$; simply set $g'$ to unity for all modes within some boundary lying just inside the diffraction limit and zero otherwise. However, this may not be a very good idea; the sharp edge of such a filter function in $k$-space will result in ringing in real space, both in the filter $g'(r)$ and especially in $H_s(r)$. These extended wings have no effect on real signal, but will couple to the incoherent noise in the images, whose spectrum does not fall to zero as one approaches the diffraction limit. To ameliorate these problems, one can choose $g'$ to have a soft rolloff as one approaches $k_{\text{max}}$ to counteract the divergence of $1/\tilde{g}$. One simple option is to set $g' = g$; i.e., to reconvolve with the PSF itself. In this case, $H_s(r)$ will be about as compact as the PSF, and we then have

$$f'_s = f_s - \gamma_s M_{sl} [g \otimes (r_1 \partial_{r_1} f_o) - g \otimes (r_1 \partial_{r_1} h) \otimes f_o], \quad (39)$$

or, equivalently,

$$f'_s = f_s + \gamma_s M_{sl} [2(r_1 \partial_{r_1} g) \otimes f_o - r_1 \partial_{r_1} g \otimes f_o], \quad (40)$$

where we have integrated by parts to avoid explicitly differentiating the image $f_o$.

To summarize, we have shown in equation (33) that the effect on the observed sky $f_o$ of a weak shear applied before seeing can be written as a shear applied after seeing plus a convolution with some kernel, which is the sheared transform of the logarithm of the OTF, which can be computed from the PSF. We have explored the form of this kernel for both Gaussian and more realistic models for the PSF. For atmospheric turbulence–limited seeing, the kernel is a power law, $H \sim r^{-11/3}$. For diffraction–limited seeing, the shear operator appears formally ill-defined, but this is not a serious problem when applied to real data; one can, for example, compute the operator for a slightly smoothed sky, $f_s = g^\prime \otimes f_o$, in a divergence-free manner. We have presented the shear operator for the case where $g^\prime = g$. This could potentially be used to compute the response of the shape statistics measured by the FOCAS and/or SExtractor packages, since these are measured from just such a reconvolved image, but we will not pursue this here.

3. WEIGHTED MOMENT SHEAR ESTIMATORS

We now specialize to weighted quadrupole moments, as defined in equation (24), and compute how these respond to shear. We first compute the response of the moments of an individual object, and then we compute the conditional mean response for a population of objects having given flux, size, etc. This will allow us to compute an optimal weight function for combining shear estimates from galaxies of different fluxes, sizes, and shapes.

3.1. Response of Weighted Moments for Individual Objects

Consider again, for illustration, the case of a Gaussian ellipsoid PSF, $g(r) \propto \exp (-r, m_{ij} r_j/2)$, and moments $q_{lm} = \int d^2r f_o g(r)r_l r_m$. From equations (34) and (24) we have $q_{lm} = d_{lm} + \delta q_{lm}$ with

$$\delta q_{lm} = -\gamma_s M_{sl} \int d^2r w_{lr} r_m (r_1 \partial_{r_1} f_o + m_{ij} r_j \partial_{r_1} r_i f_o). \quad (41)$$

Integrating by parts to replace derivatives of $f_o$ with derivatives of $w_{lr} r_m$, we find that the linear response of $q_{la} \equiv 1/2 M_{am} q_{lm}$ to a shear can be written as

$$\delta q_{la} = P_{\lambda \beta} \gamma_\beta, \quad (42)$$

with a “shear polarizability”

$$P_{\lambda \beta} = \int d^2r P_{\lambda \beta}(f_o) f_o(r), \quad (43)$$

and where

$$P_{\lambda \beta}(f_o) = \frac{1}{2} M_{am} M_{\beta \mu} [r_1 \partial_{r_1} r_1 r_m (r_1 \partial_{r_1} r_1 f_o)] \propto (\partial_{r_1} f_o). \quad (44)$$

For the special case of $w = 1$, i.e., unweighted moments, we find $P_{\lambda \beta} = \delta q_{la} (r_1 r_1 - m_{ij} r_j r_j)$, hence $\delta q_{la} = \gamma_\beta (q_{la} - m_{ij} r_j r_j) = 2\gamma_\beta (q_{la} - m_{ij} r_j r_j) = 2\gamma_\beta (q_{la} - m_{ij} r_j r_j)$, in accord with the result obtained in § 1. In this case, $P_{\lambda \beta}$ is a combination of zeroth and second moments of $f_o$.

For a general PSF and for moments measured from a filtered field, $f_o = g \otimes f_o$, as a weighted moment,

$$q_A = \frac{1}{2} M_{am} \int d^2r w_{lr} r_m (f_o r_l r_m), \quad (45)$$

we find using equation (33) that the response $\delta q_A$ can be cast in the same form, but now with

$$P_{\lambda \beta}(f_o) = \frac{1}{2} M_{am} M_{\beta \mu} [r_1 \partial_{r_1} r_1 r_m (r_1 \partial_{r_1} r_1 f_o)] - (r_1 h) \otimes g \otimes (\partial_{r_1} r_1 f_o), \quad (46)$$

where we have defined the correlation operator $\otimes$ such that $(a \otimes b, \equiv \int d^2r a(r)b(r') \delta(r - r'))$. If the moments are measured directly from the unfiltered images $f_o$, then one can replace $g'(r)$ with a Dirac $\delta$-function to obtain

$$P_{\lambda \beta}(f_o) = \frac{1}{2} M_{am} M_{\beta \mu} [r_1 \partial_{r_1} r_1 r_m (h r_1) \otimes (\partial_{r_1} r_1 f_o)], \quad (47)$$

whereas for the special case $g' = g$,

$$P_{\lambda \beta}(f_o) = \frac{1}{2} M_{am} M_{\beta \mu} [2r_1 [g \otimes \partial_{r_1} (w_{lr} r_m)] - g \otimes (r_1 \partial_{r_1} w_{lr} r_m)]. \quad (48)$$

The function $P_{\lambda \beta}(f_o)$ is shown in Figure 2 for a turbulence–limited PSF and for a Gaussian window function, $w(r)$. Note that equations (46) and (48) are well-defined continuous
functions even in the limit that the weight function becomes arbitrarily small, i.e., \( w(r) \to 0 \).

Equation (42), along with the appropriate expression for \( P_{4b}(r) \), tells us how the polarization statistic for an individual object formed from weighted quadrupole moments responds to a gravitational shear. This is an essential ingredient in calibrating the shear-polarization relation for a population of galaxies. As we shall see in the next section, there are some subtleties involved, but for now we note that if we simply average equation (42) over all the galaxies on a patch of sky, we have

\[
\langle q_x \rangle - \langle q_y \rangle = \langle P_{ab} \rangle \gamma_{ab},
\]

so an estimate of the net shear is given by

\[
\hat{\gamma}_z = \langle P_{ab} \rangle^{-1}(\langle q_y \rangle - \langle q_x \rangle),
\]

which is the generalization of equation (23) to weighted moments. The term \( \langle q_{ab} \rangle \) in equation (50) is the averaged observed polarization. The term \( \langle q_{ab} \rangle \) is the mean polarization generated by anisotropy of the PSF; we show how this can be dealt with below.

Ideally, in averaging polarizations, one should apply weight proportional to the square of the signal-to-noise ratio, which one would expect to be a function of the flux, size, eccentricity, etc., of the objects. If the shape is measured with some fairly compact window function, \( w(r) \), then the total flux, which may be dominated by the profile of the object at considerably larger radius, is probably not ideal, and one will likely obtain a better performance if one takes the weight function to be a function of \( q_0, q^2 \equiv q_x q_y \), and a weighted flux

\[
F = \int d^2r \; w(r) f_s(r).
\]

The response of \( F \) can be computed in much the same way as \( q_{4z} \), and from equation (38) we find \( \delta F = R_x \gamma_{xz} \), where

\[
R_x = \int d^2r \; R_x f_s(r),
\]

or, for the case \( g^\dagger = g \),

\[
R_x(r) = M_{ab}[\{2r(g \otimes \partial \gamma)w] - g \otimes (\partial \gamma w)\}].
\]

In the above equations we have implicitly assumed that applying a shear does not affect the location of an object. This is not necessarily the case. If objects are detected as peaks of the surface brightness, \( f_s \), smoothed with some detection filter \( w_d \), that is, as peaks of \( f_d = w_d \otimes f_s \), then for an object that in the absence of shear lies at the origin, we have \( d_i(0) = \tilde{\gamma} \mathcal{L}(w_d \otimes f_s)_0 = \int d^2r \; \tilde{\gamma} \mathcal{L}(\partial_r w_d(r) f_s(r) = 0 \), while after applying a shear we have \( d_i(0) = \gamma_{xz} M_{ab} \int d^2r \; \tilde{\gamma} \mathcal{L}(\partial_r w_d(r) \partial_r \gamma w) - (r \partial_r \gamma w) \rangle \). This will not, in general, vanish, implying that the peak location will have shifted, and consequently the central second moments should be measured about the shifted peak location, whereas in the above formulæ we have computed the change in the moments without taking the shift into account. One could incorporate this effect, but at the expense of considerable complication of the results. There is some reason to think that this effect is rather weak. In particular, if the galaxy is symmetric under rotation by \( 180^\circ \), so that \( f_s(r) = f_s(-r) \), then the shift in the centroid

\[
\begin{align*}
\text{Left: Pair of functions } w_1, w_2 \text{ with } w_i(r) = \frac{1}{2} M_{a0} w(r) r^2, \text{ which when multiplied by } f_s \text{ and integrated give the polarization statistic } q_s. \\
\text{Center and Right: Components of the polarizability kernel } \mathcal{R}_{ij}(r), \text{ which when multiplied by } f_s \text{ and integrated yields the polarizability } P_{ij}. \text{ The PSF was computed from a turbulence-limited OTF } \tilde{\gamma} = \exp \left[-0.5(r/\epsilon_0)^2 \right], \text{ and the smoothing kernel was } w(r) = \exp \left[-0.5(r/\epsilon_0)^2 \right], \text{ with scale length } \epsilon_0 \text{ equal to } \frac{1}{2} \text{ the box side.}
\end{align*}
\]
vanishes. In general this is not the case, and the formulae above should be considered only an approximation.

3.2. Response of the Population

Equation (50) above gives a properly calibrated estimate of the gravitational shear. It is, however, less than ideal as the polarization average is taken over all galaxies with equal weight. This is neither desirable nor is it achievable in practice due to selection limits; what one would rather have is an expression for the average induced polarization for all galaxies in some cell of flux, size, and shape space, which we parameterize by \( F, q_0, \) and \( q_z \). One can then average appropriately weighted combinations of the average shear estimate for each cell. The mean induced polarization for such a cell depends not only on the polarizabilities of the objects contained therein, but also on the gradient of the mean density of objects as a function of the photometric parameters \( F, q_0, \) and \( q_z \). Consider a slice through this 4-space at constant \( F \) and \( q_0 \). A shear will induce a general flow of particles in this space in the direction \( \delta a = \gamma a \). The mean polarization for a cell in \( F-q_0-q_z \) is the average around an annulus in \( q_z \) space, and depends quite sensitively on the local slope of the distribution of particles in \( |q_z| \). These factors can have a profound influence on the weighting scheme; for a distribution that is flat near the origin in \( q_z \) space, such as a Gaussian, for example, nearly circular objects have no response and should therefore receive no weight. For a randomly oriented distribution of circular disk galaxies, in contrast, the distribution in \( q_z \)-space has a cusp at the origin; the response becomes asymptotically infinite, and these objects dominate the optimally weighted combination. These examples are both idealized, but underline the importance of computing the population response in order to obtain an optimal signal-to-noise ratio.

Let us first compute the conditional mean polarization for galaxies of a given flux and size, \( \langle q_z \rangle_{F,q_0} \). The mapping of the photometric parameters \( F, q_0, \) and \( q_z \) is

\[
F' = F + \gamma_F, \quad q_0' = q_0 + \gamma_{q_0}, \quad q_z' = q_z + \gamma_{q_z},
\]

so we need to consider the distribution of galaxies in \( F, q_0, q_z, R_z, P_{os}, \) and \( P_{ab} \), with lensed and unlensed distribution functions related by

\[
n'(F, q_0, q_z, R_z, P_{os}, P_{ab}, \rho_{p_0})dF' dq_0' dR_z' dP_{os}' dP_{ab}' = n(F, q_0, q_z, R_z, P_{os}, P_{ab}, \rho_{p_0})dF dq_0 dR_z dP_{os} dP_{ab}.
\]

Multiplying by \( W(F', q_0, \rho_{p_0})q_{z'} \), where \( W \) is some arbitrary function, and integrating over all variables, we have

\[
\int dF' dq_0' dR_z' dP_{os}' dP_{ab}' n'(F' \mid F, q_0, q_z)\rho_{p_0} = \int dF dq_0 dR_z dP_{os} dP_{ab} n(F, q_0, q_z)\rho_{p_0} W(F, q_0, \rho_{p_0})q_{z'}.
\]

To zeroth order in \( \gamma \) this vanishes, because of the statistical anisotropy of the unlensed population, so by using equation (54) for \( \delta F \), etc., performing a Taylor expansion of \( W(F + \delta F, q_0 + \delta q_0) \), and integrating by parts, we have

\[
\int dF' \ldots d^4P' n'(F', q_0, q_z)\rho_{p_0} = \gamma_F \int dF \ldots d^4P W(q_0) \left( nP_{ab} - P_{os} q_s \frac{\partial n}{\partial q_0} - R_q q_s \frac{\partial n}{\partial F} \right). \quad (57)
\]

To first order in \( \gamma \), we can replace unprimed by primed quantities throughout the integral on the right-hand side, since we need to compute this only to zeroth-order accuracy. We now have a relation between the mean value of the observed polarization on the left-hand side and some other observable, again integrated over the distribution of observed galaxy properties (rather than of the unlensed parent distribution). Since \( W(q_0) \) is arbitrary, and dropping primes, this implies

\[
\gamma_F \int d^2q d^2R d^2P_0 d^4P W(q_0) \left( nP_{ab} - P_{os} q_s \frac{\partial n}{\partial q_0} - R_q q_s \frac{\partial n}{\partial F} \right)
\]

or, equivalently, that the conditional average polarization is

\[
\langle q_z \rangle_{F, q_0} = \gamma_F \left( \frac{\langle P_{ab} \rangle - \frac{1}{n} \frac{\partial n}{\partial q_0} \langle q_s \rangle}{- \frac{1}{n} \frac{\partial n}{\partial F} \langle q_s \rangle} \right), \quad (58)
\]

where \( n = n(F, q_0) \). Thus, as expected, the shear-induced shift in the mean polarization for galaxies of a given size and flux differs from the mean of the shift \( \delta q_z = P_{ab} \gamma_F \) for an individual object. This is because a shear changes the weighted flux and size of an object in a way that is correlated with its shape. When we average the shear for galaxies in some cell in flux-size space, we are averaging over galaxies that have been scattered in size and flux, and we obtain a bias in the net polarization that depends on the gradients of the distribution function.

We can generalize this to compute the response for galaxies of given flux, \( F \), size, \( q_0 \), and rotationally invariant shape parameter, \( q^2 = q_a q_s \). To do this, we set \( q_z = q_a q_s \), with the unit polarization vector \( \hat{q}_z = \langle \cos \varphi, \sin \varphi \rangle \), so \( d^2q = dq dq \varphi \). We then have, now for some arbitrary function \( W(F, q_0, q^2) \),

\[
\int dF dq_0 dq^2 d\varphi d^2R d^2P_0 d^4P n'(F, q_0, q^2)\rho_{p_0} = \int dF dq_0 dq^2 d\varphi d^2R d^2P_0 d^4P W(F, q_0, q^2)\rho_{p_0}
\]

\[
\times (F + \delta F, q_0 + \delta q_0, q^2 + \delta q^2) q_s + \delta q_s), \quad (60)
\]

where now \( n = n(F, q_0, q^2, \varphi, R_z, P_{os}, P_{ab}) \). Using \( \gamma F, \) etc., from equation (54) and equation (57), we have

\[
\int d\varphi d^2R d^2P_0 d^4P n_{q_s} = \gamma_F \int d\varphi d^2R d^2P_0 d^4P
\]

\[
\times \left( nP_{ab} - P_{os} q_s \frac{\partial n}{\partial q_0} - R_q q_s \frac{\partial n}{\partial F} - 2P_{ab} \frac{\partial n q_s}{\partial q^2} \right), \quad (61)
\]
or, equivalently, \( \langle q_x \rangle_{F, q_0, q^2} = \bar{P}_{sb} \gamma \), with effective polarizability

\[
\bar{P}_{sb} = \langle P_{sb} \rangle - \frac{2}{n} \frac{\partial n}{\partial q^2} \frac{\langle q_x P_{sb} q_x \rangle}{\langle q_x \rangle_{F, q_0, q^2}} - \frac{1}{n} \frac{\partial n}{\partial q_0} - \frac{1}{n} \frac{\partial n}{\partial F} \left( \frac{\partial}{\partial \ln q} - \frac{2}{\partial \ln q^2} \right) \langle R_{sb} q_x \rangle,
\]

where now \( n = n(F, q_0, q^2) \), and all averages are at fixed \( F, q_0, \) and \( q^2 \).

The photometric parameters \( q_0, q^2 \) appearing here are unnormalized. This is convenient for computing the linear response functions. It is, however, somewhat awkward here, since the distribution function \( n(F, q_0, q^2) \) is highly skewed, because \( q_0 \) and \( q^2 \) correlate very strongly with the flux. In computing the effective polarizability, it is more convenient to work with rescaled variables, \( q_0' = q_0/F \) and \( q^2' = q^2/F^2 \). The distribution function in rescaled variables is

\[
n'(F, q_0', q^2') = F^3 n(F, q_0, q^2) .
\]

If we also rescale the polarizabilities, \( R' = \frac{R}{F}, \bar{P}_{sb}' = \frac{P_{sb}}{F}, P_{sb} = \bar{P}_{sb} F^2 \), reexpress equation (62) entirely in terms of primed quantities, and then drop the primes, we find

\[
\bar{P}_{sb} = \langle P_{sb} \rangle - \frac{2}{n} \frac{\partial n}{\partial q^2} \frac{\langle q_x P_{sb} q_x \rangle}{\langle q_x \rangle_{F, q_0, q^2}} - \frac{1}{n} \frac{\partial n}{\partial q_0} - \frac{1}{n} \frac{\partial n}{\partial F} \left( \frac{\partial}{\partial \ln q} + \frac{\partial}{\partial \ln q_0} + \frac{\partial}{\partial \ln q^2} \right) \langle R_{sb} q_x \rangle,
\]

where we have used the result that for any function \( X(F, q_0, q^2) \), the partial derivative with respect to \( F \) at constant \( q_0, q^2 \) is \( \frac{\partial X(F, q_0, q^2)}{\partial F} = \frac{\partial}{\partial \ln F} X + \frac{q^2}{\partial \ln q^2} X \) to compute the term involving \( R_s \). The rather cumbersome equation (64) calibrates the relation between the shear and the mean polarization for galaxies in a small cell in flux-size-shape space. To compute it, we need to bin galaxies in this space to obtain the mean density \( n \) and the various averages appearing here, and then perform the indicated partial differentiation. The form of equation (64) is somewhat inconvenient, since the density \( n(F, q_0, q^2) \) is asymptotically constant as \( q^2 \to 0 \), and one must properly deal with the discontinuous derivative at this boundary. A computationally more convenient approach is to make one final transformation from \( q^2 \to q \), since \( n(F, q, q^2) = 2q n(F, q_0, q^2) \) falls to zero as \( q \to 0 \), and there is no need for any special treatment of the derivatives at the boundary. With this transformation, we have

\[
\bar{P}_{sb} = \langle P_{sb} \rangle - \frac{1}{n} \frac{\partial n}{\partial q} \frac{\langle q_x P_{sb} q_x \rangle}{\langle q_x \rangle_{F, q_0, q^2}} - \frac{1}{n} \frac{\partial n}{\partial q_0} \frac{\langle R_{sb} q_x \rangle}{\langle q_x \rangle_{F, q_0, q^2}} + \frac{q}{n} \left( \frac{\partial}{\partial \ln F} + \frac{\partial}{\partial \ln q_0} + \frac{\partial}{\partial \ln q^2} \right) \langle R_{sb} q_x \rangle,
\]

where now \( n = n(F, q_0, q) \).

What about measurement noise? Let us assume that one has been given an image containing signal and measurement noise, and that one has detected objects, and measured quantities such as \( F \) and \( q_A \). How would these photometric parameters change under the influence of a gravitational shear? The answer is given by equation (54), but with the understanding that \( R_{sb}, P_{sb} \), etc., are the response functions one would measure in the absence of noise. This means that equation (65) is also applicable, with the same proviso. The major terms in equation (65) are, however, invariant of additive noise. The exceptions are the terms involving \( \langle P_{sb} q_x \rangle \) and \( \langle R_{sb} q_x \rangle \), which are quadratic in the sky surface brightness. The measured expectation values \( \langle P_{sb} q_x \rangle \), etc., therefore exceed the true values, but by an amount that one can calculate from the known properties of the measurement noise. Another implicit assumption in the above analysis is that the objects are actually detected, which restricts applicability to objects that are detected at a reasonable level of significance. Aside from this, the results above should be applicable in the presence of measurement noise. To test these claims, we have made extensive simulations with mock data, the details of which are described in Appendix B. Figure 3 shows the results of one of these. The actual polarization agrees quite closely with the effective polarizability, even for very faint objects. While the differences between the effective polarizability and that for individual objects is not very large (typically on the order of 20% or so), the effective polarizability clearly describes the true response more faithfully. We now use equation (65) to construct a minimum variance weighting scheme for combining shear estimates.

### 3.3. Optimal Weighting with Flux, Size, and Shape

Armed with the conditional mean polarization \( \langle q_x \rangle_{F, q_0, q^2} \), we can now compute an optimal weight (as a function of, for example, flux \( F \), size \( q_0 \), and eccentricity \( q^2 \) ) for combining the estimates of the shear from galaxies of different types. Let us assume that one has measured fluxes, etc., for a very large number of galaxies—from an entire survey, say—and that from these data one has determined the mean number density of galaxies, \( n(F, q_0, q^2) \), and also the various conditional averages and gradients that appear in equation (65). Now consider a relatively small spatial subsample of these galaxies and bin these into cells in \( (F, q_0, q^2) \) space, and for each bin compute the occupation number, \( N \), and the summed polarization, \( \sum q_x \).

A shear estimator for a cell that happens to have occupation number \( N \) is

\[
\hat{\gamma} = \bar{P}_{sb}^{-1} \sum q_x / N .
\]

To simplify matters, let us neglect for now any anisotropy of the PSF, in which case we can write \( \bar{P}_{sb} = \bar{P}_{sb} \), where \( \bar{P} = \bar{P}_{sb} / 2 \); we then have

\[
\hat{\gamma} = \frac{1}{N \bar{P}} \sum q_x 
\]

The expectation value for the variance in \( \hat{\gamma} \) (for a cell that happens to contain \( N \) galaxies) is

\[
\langle \hat{\gamma}^2 \rangle = \frac{\langle \sum q_x \rangle^2}{N^2 \bar{P}^2} = \frac{q^2}{N \bar{P}^2} ,
\]

where we have used \( \langle \sum q_x \sum q_x \rangle = \langle \sum q^2 \rangle \), since in the limit of weak shear the galaxy polarizations are uncor-
related. Since different cells give shear estimates whose fluctuations are mutually uncorrelated, the optimal way to combine the shear estimates from all the cells is to average them with weight per cell, \( W_{\text{cell}} \propto 1/\langle \hat{\gamma}^2 \rangle = N P^2/q^2 \), to obtain a final optimized total shear estimate,

\[
\hat{\gamma}^\text{total}_{\gamma} = \frac{\sum_{\text{cells}} (N P^2/q^2)[(\sum q_{\text{cell}})/N P]}{\sum_{\text{cells}} N P^2/q^2} = \frac{\sum_{\text{galaxies}} P q_{\text{galaxy}}/q^2}{\sum_{\text{galaxies}} P^2/q^2} = \frac{\sum_{\text{galaxies}} Q q_{\text{galaxy}}}{\sum_{\text{galaxies}} Q^2},
\]

where \( Q \equiv P/q \). Thus, the optimized cell weighting scheme corresponds to averaging the shear estimates from individual galaxies, \( \hat{\gamma}_{\gamma}^\text{galaxy} = q/P \) (for galaxies of given \( F, q_0, q^2 \)), with weights \( W_{\text{galaxy}} = P^2/q^2 \).

The variance in the total shear estimator is

\[
\langle \gamma^2 \rangle = \frac{\sum Q^2 \langle \hat{\gamma}^2 \rangle}{(\sum Q^2)^2} = (\sum Q^2)^{-1}. \tag{70}
\]

The quantity \( \sum Q^2 \) is extensive with the number of galaxies, and its value, per unit solid angle of sky, provides a useful figure of merit for weak-lensing data. From a 2.75 hr \( I \)-band integrations of solid angle \( d\Omega = 0.165 \text{ deg}^2 \) at the Canada-France-Hawaii Telescope (CFHT) taken in good seeing (0.60 FWHM), we obtained \( \sum Q^2 \approx 4.7 \times 10^4 \) (Kaiser et al. 1998) or \( \sum Q^2/d\Omega \approx 2.85 \times 10^5 \text{ deg}^{-2} \); with data of this
quality, the statistical uncertainty in the net shear (per component) measured over 1 deg$^2$ would be around $\sigma_r \approx (2 \times 2.85 \times 10^{-3})^{-1/2} = 1.32 \times 10^{-3}$.

This figure of merit allows one to tune the parameters of one’s shape-measurement scheme, such as the weight function scale size, in an unbiased and objective manner. The weighting scheme derived above is appropriate if the shear is independent of the measured flux, etc. This is the case for lensing by low-redshift clusters, where for all relevant values of source redshift the sources are effectively “at infinity,” and the shear has saturated at its value for an infinitely distant source, $\gamma_\infty$. For high-redshift lenses, the shear will vary with source redshift, and if one has some, perhaps probabilistic, distance information at one’s disposal, then the weighting scheme should be modified. Let us assume that the measured photometric properties, $p_i$, of some object indicate that it has a probability distribution to be at distance $z$ of $p(z|p_i)$; with high-resolution spectroscopy this would be a $\delta$-function at the measured redshift, whereas with only broadband colors the conditional probability would be smeared out and perhaps multimodal. The conditional mean shear for this object, for a given a foreground lens, is proportional to the mean inverse critical surface density, and the optimal estimate for the shear at infinity is

$$\tilde{\gamma}_i = \frac{1}{\Sigma_{\text{crit}}(\infty)} \sum_{\text{galaxies}} \langle 1/\Sigma_{\text{crit}}(z) \rangle Q q_{i}.$$  

4. CORRECTION FOR PSF ANISOTROPY

In the foregoing discussion we have computed how the shape polarization $q_{i}$ responds to a shear. This allows us to correctly calculate the circularizing effect of seeing. In general, asymmetry of the PSF will also introduce a spurious systematic shape polarization, $\langle q_{i} \rangle$, in equation (50), and it is crucial that this be measured and corrected for. As discussed in § 1, for unweighted second moments, the effect of PSF anisotropy is rather simple, since the (flux-normalized) final second moment $q_{i\mu}$ is the sum of the intrinsic second moment and that of the instrument, so the anisotropic parts of the instrumental second moment, $q_{i\mu} = M_{i\mu} q_{i\mu}$, can be measured from stars and subtracted from those observed. For weighted or isophotal moments things are more complicated, and depend on how the anisotropy is generated. KSB considered a simple model in which the final PSF is the convolution of some circularly symmetric PSF with a small, but highly asymmetric, “anisotropizing kernel,” $k(r)$. In § 4.1 we will further explore this “convolution model” and the attempts that have been made to implement this perturbative approach, and estimate the error in this method. In § 4.2 we show that the convolution model is quite inappropriate when applied to diffraction-limited seeing, and we develop a more general technique for correcting PSF anisotropy. Finally, in § 4.3 we draw attention to a type of noise-related bias that affects shear estimators, but which has hitherto been overlooked, and we show how this can be dealt with.

The artificial polarization produced by PSF anisotropy is an effect that is present in the absence of any gravitational shear. Thus, at leading order in $\gamma$ we can set $\gamma = 0$. This effectively decouples the computation and correction of the PSF anisotropy from the shear polarization calibration problem considered above. It also means that in this section we can assume that the intrinsic shapes of galaxies are statistically isotropic.

4.1. The Convolution Model

KSB computed the effect of a PSF anisotropy under the assumption that this can be modeled as the convolution of a perfect circular PSF with some compact kernel $k(r)$. This would be a good model, for instance, for observations primarily limited by atmospheric turbulence, with seeing disk of radius $r_\text{w}$, but with very small guiding or registrations errors or optical aberrations with extent $\delta r \ll r_\text{w}$, and the KSB analysis gives a correction to lowest order in $\delta r$. In this model, the effect of “switching on” the anisotropy is the transformation on the observed image, which is smooth on scales $\sim r_\text{w}$,

$$f_o(r) \rightarrow f_a(r) = \int d^2 r' k(p(r')) f_o(r - r')$$

$$= f_o(r) - k_i \hat{\partial}_i f_o(r) + \frac{1}{2} k_{ij} \hat{\partial}_i \hat{\partial}_j f_o(r)$$

$$- \frac{1}{2} k_{ijk} \hat{\partial}_i \hat{\partial}_j \hat{\partial}_k f_o(r) + \ldots,$$

(72)

where we have Taylor expanded $f_o(r - r')$ and where $k_i \equiv \int d^2 r r_i k(r)$, $k_{ij} \equiv \int d^2 r r_i r_j k(r)$, etc. Taking the kernel to be centered, we have $k_i = 0$, and to lowest nonvanishing order the effect on the polarization is

$$q_a = \int d^2 r w_a(r) f_a(r) = q_a + \frac{1}{2} k_{ij} \int d^2 r f_o(r) \hat{\partial}_i \hat{\partial}_j w_a(r),$$

(73)

where $w_a \equiv M_{\text{alm}} w(r_1 r_2)$, and we have integrated by parts. The induced stellar polarization is linear in $k_i$, and therefore scales as the square of the extent of the convolving kernel, since $k_i \sim \delta r^2$. Performing the decomposition $k_{ij} = k_{ij} M_{\text{alm}}$, we find that an individual galaxy’s polarization, $q_\mu$, will depend on both the trace $k_0$ and the trace-free parts $k_{ij}$ of $k_{ij}$. The average induced polarization, however, only depends on $k_0$, and we have $\langle q_\mu \rangle = k_0 P_{\text{alm}}$, with

$$P_{\text{alm}} = \frac{1}{2} M_{\text{alm}} M_{\text{alm}} \int d^2 r \langle f_o(r) \hat{\partial}_i \hat{\partial}_j w(r_1 r_2) \rangle.$$  

(74)

To lowest order in the PSF anisotropy, we can use the KSB polarizability equation (74) to infer $q_\mu$ from the shapes of stars, and then correct the weighted second moments of the galaxies. At the same level of precision, one can convolve one’s image with a small kernel designed to nullify the anisotropy. Fischer & Tyson (1997) have presented a $3 \times 3$ smoothing kernel that does this. In the typical situation, the shapes are measured from an average of numerous images taken with a pattern of shifts, which are coregistered to fractional pixel precision; a simpler but equally effective approach is then to average over pairs of images that have been deliberately displaced from the true solution by the small displacement vector $\pm \delta r_\text{w} = |k_{ij}|^{1/2} \cos(\theta) \sin(\theta)$, with $\theta = \tan^{-1} (k_{ij}/k_i) + \pi/2$.

What is the error in this linearized approximation? To answer this, we must consider higher terms in the expansion equation (72). The next-order correction to $q_\mu$ involves $k_{ijk}$, which unlike the centroid $k_0$ cannot be set to zero, so for a given object the fractional error in the KSB correction is on the order of $\delta r / r_\text{w} \sim e^{1/2}$. Galaxies, however, are randomly oriented on the sky, so the average change in $q_\mu$ for a galaxy of some arbitrary morphology but averaged over all posi-
tion angles at this order in \( \delta r/r_0 \) is

\[
\langle \delta q_x \rangle \sim k_{ijk} \int d^2 r \langle f_\delta(r) \rangle \partial_i \partial_j \partial_k w_x(r). \tag{75}
\]

This vanishes, since \( \langle f_\delta(r) \rangle \) is an even function while \( \partial_i \partial_j \partial_k w_x(r) \) is odd, provided that we take \( w(r) \) to be circularly symmetric at least, which we will assume is the case. The effective net fractional error in the KSB approximation is therefore on the order of \( (\delta r/r_0)^2 \), or typically on the order of the induced stellar ellipticity.

### 4.2. General PSF Anisotropy Correction

In principle, one can develop a higher order correction scheme within the context of this model, but this does not seem to be particularly promising. It is not clear that the result will be sufficiently robust and accurate even for ground-based observations; in very good seeing conditions the PSF anisotropy from aberrations becomes large, and the perturbative approach will break down. Moreover, for observations with telescopes in space, the model of the PSF as a convolution of a perfect circular PSF with a kernel, small or otherwise, is wholly unfounded. The instantaneous OTF is

\[
\tilde{g}(k) = \int d^2 r A(r) A\left( r + \frac{k L \lambda}{2 \pi} \right) \exp \left\{ \left[ \phi(r) - \phi\left( r + \frac{k L \lambda}{2 \pi} \right) \right] \right\}.
\tag{76}
\]

where \( A \) is the real transmission function of the telescope input pupil, \( L \) is the focal length, and \( \phi(r) \) is the phase error due to mirror aberrations and the atmosphere. It is often said that the general OTF factorizes into a set of terms describing the atmosphere, the telescope aperture, and aberrations, and this would seem to justify the convolution model discussed above. For atmospheric turbulence, and for aberrations arising from random small-scale mirror roughness, this is correct. This is because the average of the complex exponential term in equation (76) depends only on \( \delta r = k L \lambda/2 \pi \) and is independent of \( r \); equation (76) then factorizes into two independent terms, and the PSF is the convolution of two completely independent and non-negative functions, but this is not the case in general. To be sure, one can write the combined aperture and aberration OTF as a product of some “perfect” OTF with some other function (the true OTF divided by the perfect one), but quite unlike the case for random turbulence and mirror roughness, this function is neither independent of the shape of the pupil nor is it positive; if you compute this function for a telescope subject to a low-order aberration from figure error, for instance, you will find that the this function is strongly oscillating and is just as extended as the true PSF. Consider the situation in WFPC2 observations, for example, where the PSF anisotropy has important contributions both from asymmetry of the pupil, \( A(r) \), and from phase errors, although the former tends to dominate (although not enormously so) for long wavelengths and far off-axis with the WFPC2. The aperture function for the lower left corner of chip 2 computed from the Tiny Tim model (Krist 1995) is shown in Figure 4. The off-axis pupil function is approximately that on-axis minus a disk of some radius \( r_1 \) at some distance \( \Delta r \) off-axis. If we let the radius of the primary be \( r_0 \) and define a disk function \( D_{\Delta r}(r) = \Theta(r/r_0 - 1) \), then on computing the electric field amplitude, \( a(x) \), and squaring, we find that the on-axis PSF is of course given by the Airy disk, \( g \approx D_0 \), while the off-axis PSF \( g' \) is given by \( g'(x) \approx g(x) - D_{\Delta r} \cos(2 \pi x \Delta r/\lambda L) \). For \( r_1 \ll r_0 \), \( D_{\Delta r} \) is relatively slowly varying and close to unity, in which case the extra off-axis obscuration introduces a pertur-

**Fig. 4.—**Left: WFPC2 pupil function and phase error from Krist’s Tiny Tim model for an off-axis point on the focal plane. The phase is given in radians for a wavelength of 800 nm. Right: Modulation transfer function (the real part of the OTF).

**Fig. 5.—**HST WFPC2 PSF from Tiny Tim and recircularizing filter computed as described in the text.
bation that is proportional to the product of $D_0$ and a planar wave, or essentially an asymmetric modulation of the side lobes of the on-axis PSF (see Fig. 5, left).

In general, then, and especially for diffraction-limited observations, the perturbative convolution model cannot be trusted. Luckily, at least in the context of the shear estimators discussed here, a simple solution clearly presents itself: since to compute the polarizability requires that one generate a rather detailed model of the PSF, and since we probably want to apply some kind of reconvolution, $f_i \rightarrow f_i = g^\dagger \otimes f_i$, to obtain a well-behaved shear response, we might as well use the opportunity to choose $g(r)$ in order to recircularize the PSF exactly. For example, from the observed PSF, one can compute the OTF $\hat{g}(k)$ and then form the circularly symmetric function $\hat{g}_{\min}(k)$, that being the greatest real function that lies everywhere below $|\hat{g}(k)|$. The function $g(r)$ with transform

$$g^\dagger(k) = \frac{\hat{g}_{\min}(k)^2}{\hat{g}(k)}$$  (77)

both guarantees a well-defined shear polarizability and gives a perfectly circular resulting total PSF. This is illustrated in Figure 5.

In some situations, the PSF may have near 180° symmetry, in which case a good approximation is to reconvolve with a PSF that has been rotated through 90°. Since the PSF is real, the OTF must satisfy the symmetry $\hat{g}(k) = \tilde{g}^*(−k)$. In the absence of aberrations, a diffraction-limited telescope has OTF $\hat{g} = A \otimes A$, which is real and so has exact symmetry under rotation by π, a property that is shared by the PSF, so $R_\pi g = g$. If one convolves with $g' = R_{\pi/2} g$, then the resulting total PSF, $g^\dagger \otimes g'$, is symmetric under 90° rotations (proof: $R_{\pi/2} \hat{g}_{\min} = R_{\pi/2}(R_{\pi/2} \hat{g}) = R_\pi \hat{g} R_{\pi/2} \hat{g} = \tilde{g} \hat{g}_{\min} = \hat{g}_{\min}$). For a galaxy of a given form, the average PSF-induced polarization is

$$\langle q_x \rangle = \int d^2 r \langle \hat{g}_{\min} \otimes f \rangle w_z = \langle \hat{g}_{\min} \otimes f \rangle w_z |_{r=0} = \int d^2 k \langle \tilde{f} \rangle \hat{g}_{\min} \tilde{w}_z,$$  (78)

where $w_z = M_{z,m} w(r) r_{m}^2$. But $\langle \tilde{f} \rangle$ is circularly symmetric and $\tilde{w}_z$ is antisymmetric under 90° rotations, so $\langle q_x \rangle = 0$. Other situations in which this would work would be for ground-based observations with large-amplitude telescope oscillations, and in fast guiding. In general, however, aberrations from figure errors will introduce both real and imaginary contributions to the OTF, and the latter will destroy the exact 180° symmetry of the PSF. This is certainly the case for the WFCPC2 PSF at optical wavelengths.

In the linearized approximate schemes described earlier, the PSF anisotropy is entirely characterized by the two coefficients $k_x$. In general, these will vary substantially with position on the image, but this can be treated by modeling $k_x$ as some smoothly varying function of image position $R$, such as a low-order polynomial. In the scheme proposed here, we need to be able to generate not just these two coefficients, but the full two-dimensional PSF $g(r, R)$ (being the intensity at distance $r$ from the centroid of a star at position $R$). A simple practical approach is to solve for a model $g(r, R) = \sum_{n} g_n(r)f_n(R)$, where $f_n(R)$ is some set of polynomial or other basis functions. A least-squares fit for the image-valued mode coefficients is $g_n(r) = m_{ij}^{-1} G_{ij}(r)$, with $m_{ij} = \sum_{stars} f_j(r)f_j(r)$; $G_{ij}(r) = \sum_{stars} f_j(r)g_{\min}(r)$. Armed with the image-valued coefficients, one can then, for example, compute the convolution $f_i = g \otimes f_i$, as $f_i(r) = \sum_{j} f_j(r)g_{ij}(r) \otimes f_j$; that is, we convolve the source image with each of components $g_i(r)$, and then combine with spatially varying coefficients $f_j(R)$. Nonlinear functions of the PSF, such as the recircularizing kernel $g^\dagger$, are easily generated by making realizations of the PSF models on, say, a coarse grid of points covering the source image, and from each of these computing the desired function, then fitting the results to a low-order polynomial, just as for $g(r, R)$.

The approach described here lends itself very nicely to observations with large mosaic CCD cameras, in which misalignment of the chip surfaces and the focal plane, coupled with telescope aberrations, can give rise to a PSF that varies smoothly across each chip, but changes discontinuously across the chip boundaries. This results in a very complicated PSF pattern on the final image, obtained by averaging over many dithered images. It is a good deal simpler to recircularize each of the contributing images. We still need to generate a spatially varying model of the final $g(r)$ and $g^\dagger(r)$ in order to compute the polarizability, but if we fail to model these exactly, it will only introduce a relatively minor error in the shear polarization calibration.

4.3. Noise Bias

In the absence of noise, the recircularization procedure exactly annihilates any PSF anisotropy; the signal content of the image is exactly statistically isotropic. The noise component in the images (which in the original images is incoherent Poisson noise) will, however, be correlated with an anisotropic two-point function: the peaks and troughs of the noise will appear as ellipses with correlated position angles. For a given object, the noise is equally likely to produce a positive or negative fluctuation in the polarization $q_x$, and the effects cancel out on average. Due to the nature of faint galaxy counts, however, objects of a given observed flux are more likely to be intrinsically fainter objects that have been scattered upward in flux than intrinsically brighter objects that have been scattered down, and there will therefore be a tendency for faint objects to be aligned like the recircularizing PSF; i.e., oriented opposite to the PSF. We will refer to this as “noise bias.” It should noted that an analogous effect is present in the old KSB anisotropy correction scheme. In that case the noise is isotropic, but objects are detected as peaks of a smoothed image, and there is then a tendency for the galaxies close to the detection threshold to be aligned like the PSF. In the context of perturbative schemes as described in § 4.1, there are other uncalculated errors in the correction that are typically of similar order as the noise bias effect, which goes some way to explain why it has been ignored in the past. In the improved PSF anisotropy correction method proposed here, it is the leading source of error, and it behooves us to analyze and correct for it.

Noise will affect the photometric parameters such as $q_x$ in two ways: in addition to a straightforward additive noise term, noise will also affect the object detection, and will shift, e.g., the centroid about which we measure the centered second moments. In fact, the decentering effect is quite weak. To see why, consider the following simple (although actually quite practical and realistic) model for the detection and measurement process, in which we take the recir-
cularized image $f_s$ and define objects as peaks of the field $F(r) = w \otimes f_s$, where $w$ is a some smooth, circularly symmetric weight function (we typically use a compensated “Mexican hat” filter), and the object moments, $q_A$, are the value of the field, $q_A(r) = w_A \otimes f_s$, evaluated at the peak location. Let us assume that there exists some object which, in the absence of noise, would lie at the origin; for this object we would have

$$F = F(0) = \int d^2 r w(r) f_s(r) ,$$

$$0 = d_A(0) = \int d^2 r (\partial_i w) f_s(r) ,$$

$$q_A = q_A(0) = \int d^2 r w_A(r) f_s(r) ,$$

(79)

where the condition that the object be a peak of $F$ is expressed as the vanishing of the dipole-like quantity $d_A(r) = \partial_i F$. If we now add noise, and let $f_s \rightarrow f_s + f_n$, where $f_n(r)$ is the noise component of the recircularized image, all of the photometric fields $F(r)$, $d_A(r)$, $q_A(r)$ will change to $F^\prime = F + \delta F = F + w \otimes f_n$, etc. The dipole $d_A$ will no longer vanish at the origin, but will have some value $d_A(0) = \int d^2 r (\partial_i w) f_n$, and the peak of $F(r)$ will have moved to some position $r_{pk} \neq 0$. In the vicinity of the peak, we have $d_A(r) \approx 0 + (r - r_{pk}) \partial_i d_A(r)$, and hence, to first order in the noise amplitude, the peak location is

$$r_{pk} = - (\partial_j d_A)^{-1} \delta d_A = - \left( \int d^2 r \partial_i \partial_j w f_s \right)^{-1} \int d^2 r \partial_i \partial_j w f_s .$$

(80)

The shift of the peak has no effect on $F$ at first order (because $F$ is stationary), but it does affect the central moments, $q_A$, and we have

$$\delta F = \int d^2 r w(r) f_s(r) ,$$

$$\delta q_A = \int d^2 r [w_A(r) - v_{Ai} \partial_i w] f_n ,$$

(81)

where

$$v_{Ai} = \left[ \int d^2 r (\partial_i \partial_j w) f_s \right]^{-1} \int d^2 r (\partial_i \partial_j w_A) f_s .$$

(82)

Thus, in general, the first-order change in the second moments $q_A$ depends not just on the noise and the kernel, $w_A$, but on the form of the underlying noise-free image, $f_s$, through the term involving $v_{Ai}$. There is good reason to think that this form dependence is weak for real objects; the function $\partial_i w_A$ is an odd function, so the extra term vanishes if the galaxy is symmetric under $180^\circ$ rotations. In addition, it is difficult to see why the presence of this term would cause a systematic polarization; the change in the polarization, $\delta q_A$, associated with the centroid shift is proportional to the vector $v_{Ai}$, which, being a function of the recircularized image $f_n$, is equally likely to be positive or negative. For each galaxy that, for a given realization of noise, suffers a certain $\delta q_A$, there is a $180^\circ$ rotated clone that has precisely the same weighted flux, polarization, etc. (these being even functions), but suffers a $\delta q_A$ of opposite sign. The kind of effect envisioned in the leading paragraph of this subsection arises if there is a correlation between fluctuations in the polarization and the flux $F$, coupled with a gradient of the density of objects, $n(F)$. The first term in $\delta q_A$ in equation (81) does indeed correlate with $\delta F$, as we shall see, but when we average over the orientation of the underlying noise-free galaxies, the second does not. Unfortunately, we have not been able to rigorously demonstrate that the centroid shift term vanishes. Nonetheless, in what follows we will assume that this term is negligible. This should certainly be adequate to estimate the magnitude of the effect, and is probably sufficiently accurate to give a useful correction, although the results should strictly be regarded as only approximate.

Let us now assume that one has measured a set of $n$ photometric parameters $p_i$ for a galaxy that, like $F$, $q_0$, and $q_a$, are linear functions of the brightness, $f_n$, and of the form $p_i = \int d^2 r K_i(r) f_s(r)$. The effect of noise in the images will be to introduce a perturbation in these parameters, $\delta p_i$. These perturbations have zero mean, $\langle \delta p_i \rangle = 0$, and since there are typically a very large number of photons (noise plus signal) contributing to the parameters $p_i$, the central limit theorem dictates that the $\delta p_i$ will have a multivariate Gaussian distribution:

$$\sigma(\delta p) d^p = \left[ (2\pi)^n | C \right]^{-1/2} \exp \left( - \frac{\delta p C^{-1} \delta p}{2} \right) d^p ,$$

(83)

where the covariance matrix is $C_{ij} = \langle \delta p_i \delta p_j \rangle$ and is given by

$$C_{ij} = \int d^2 r K_i(r) K_j(r) ,$$

(84)

where $K_i(r) = \int f_s(r') f_s(r+r')$ is the two-point function of the noise $f_n(r)$, and can be computed as described in Appendix A.

For objects that are detected at a fairly high level of significance, $v \gg 1$, the noise will cause a small modification of the underlying distribution function: $n(p) \rightarrow n(p)$, which we can compute in a perturbative manner:

$$n(p) = \int d^p \delta p n(p - \delta p) \sigma(\delta p)$$

$$= n(p) - \frac{\partial n}{\partial \delta p} \delta p \left. + \frac{\partial^2 n}{2 \partial \delta p \partial \delta p} \right. + \ldots .$$

(85)

Since $\sigma(p)$ is an even function, all the odd terms in this expansion vanish, and we have $n(p) = n(p) + \frac{1}{2} C_{ij} \partial^2 n / \partial \delta p_i \partial \delta p_j$ at leading order.

Let us specialize now to the $n = 4$ dimensional case, $p_i = F, q_0, q_s$, and compute the mean polarization for galaxies of given $F$ and $q_0$ with some weight function $W(q = |q_s|)$:

$$\langle q_s \rangle_{F, q_0} = \frac{\int d^2 q W(q) n(F, q_0, q_s)}{\int d^2 q W(q) n(F, q_0, q_s)}$$

$$\approx \frac{\int d^2 q W(q) \delta n(F, q_0, q_s)}{\int d^2 q W(q) n(F, q_0, q_s)} ,$$

(86)

where $\delta n(F, q_0, q_s) = \frac{1}{2} C_{ij} \partial^2 n / \partial \delta p_i \partial \delta p_j$. Since $q_s$ is an odd function and $W(q)$ and $n(F, q_0, q_s)$ are even, the only terms
that contribute to the integral in the numerator are those involving a single derivative with respect to polarization, \( \partial^2 n / \partial F \partial q_{\delta} \) and \( \partial^2 n / \partial q_{\delta} \partial q_{\phi} \); on integrating by parts, we have

\[
\langle q_{\delta} \rangle = - \frac{1}{2} \int d^2 q \left( \frac{\partial F \partial q_{\delta}}{\partial F} \frac{\partial n}{\partial q_{\delta}} + \frac{\partial \delta n}{\partial q_{\delta}} \frac{\partial n}{\partial q_{\delta}} \right)
\]

\[
\times \frac{\partial W(q)q_s}{\partial q_{\delta}} \left[ \int d^2 q \ W(q)n(F, q_0, q_{\delta}) \right]^{-1}.
\]

Using \( \partial W(q)q_s / \partial q_{\delta} = \delta_{q_s} W(q) + q_s q_{\delta} dW / dq_{\delta} \), integrating over angle, and converting from \( n(F, q_0, q) \) to \( n'(F, q_0, q) = 2\pi n(F, q_0, q) \) and dropping the prime, we have

\[
\langle q_{\delta} \rangle = - \frac{1}{2} \int d^2 q \left( \frac{\partial F \partial q_{\delta}}{\partial F} \frac{\partial n}{\partial q_{\delta}} + \frac{\partial \delta n}{\partial q_{\delta}} \frac{\partial n}{\partial q_{\delta}} \right)
\]

\[
\times \left[ W(q) + \frac{1}{2} qW'(q) \right] \int dq nW(q) \right]^{-1},
\]

and letting \( W(q) \) become a \( \delta \)-function to isolate a single value of \( q_{\delta} \), we have

\[
n\langle q_{\delta} \rangle = - \frac{1}{2} \left( \frac{\partial F \partial q_{\delta}}{\partial F} \frac{\partial n}{\partial q_{\delta}} + \frac{\partial \delta n}{\partial q_{\delta}} \frac{\partial n}{\partial q_{\delta}} \right).
\]

Thus, as anticipated, there is a net induced polarization if there is a nonzero correlation between the polarization fluctuation, \( \delta q_{\delta} \), and \( F \) and/or \( q_0 \), and also significant gradients of the distribution function with respect to \( F \) or \( q_0 \).

The effect of the first term in equation (89) on the rescaled polarization is \( \langle q_{\delta} \rangle = \langle q_{\delta} \rangle / F = - \left( \frac{1}{2} \langle \delta F \delta q_{\delta} \rangle / F^2 \right) \ln n / \ln F \), and is second order in the inverse significance: \( \langle q_{\delta} \rangle \propto v^{-2} \), where \( v^2 = F^2 / \langle \delta F^2 \rangle \), and rapidly becomes small for well-observed objects. Thus, it should be possible to set a sensible limit on the significance of object detection, and, if necessary, to use equation (89) together with equation (84) for \( \langle \delta F \delta q_{\delta} \rangle \), etc., to correct for this. The error in this linearized approximation is of fourth order in the inverse significance.

As a simple illustrative example, consider a Gaussian galaxy of scale \( r_0 \), a Gaussian window function \( w \) with scale \( r_w \), and a Gaussian ellipsoid PSF \( g = \exp \left[ - (x^2/r_w^2 + y^2/r_w^2)/2 \right] \) with \( r_w = r_0 (1 + \epsilon / 2) \) and \( r_s = r_0 (1 - \epsilon / 2) \). A suitable recircularizing kernel is then \( g^s = \exp \left[ - (x^2/r_s^2 + y^2/r_s^2)/2 \right] \), and the normalized two-point function of the noise is

\[
\zeta_n(r) = \frac{1}{\pi r_s r_s} \exp \left( - \frac{x^2/2r_s^2 + y^2/2r_s^2}{2} \right).
\]

The expectation averages are, to first order in \( \epsilon \),

\[
\langle \delta F^2 \rangle = 2(r_w^2 + r_s^2),
\]

\[
\langle \delta q_1 \delta F \rangle = r_w^2 (r_w^2 - r_s^2) / (r_w^2 + r_s^2) = 2r_w^2 r_s^2 (r_s^2 + r_w^2)^2 / (r_s^2 + r_w^2),
\]

and, taking only the first term in equation (89) for simplicity,

\[
\langle \delta q_1 \rangle = \frac{\delta \ln n}{\delta F} = - \frac{1}{2} \frac{\delta \ln n}{\delta F} \frac{1}{v^2} \langle \delta F^2 \rangle \approx \frac{1}{v^2} \frac{r_w^2 r_s^2}{r_w^2 + r_s^2},
\]

using \( \delta \ln n / \delta F \approx 1 \), as observed. A shear of strength \( \gamma \) applied to the Gaussian galaxy produces a polarization

\[
\langle \delta q_1 \rangle = \frac{4\gamma r_w^2 r_s^2}{r_w^2 + r_s^2 + r_w^2};
\]

therefore, a PSF anisotropy of strength \( \epsilon \) is equivalent to an effective shear

\[
\gamma = \frac{\epsilon}{8v^2} \frac{r_w^2 (r_w^2 + r_s^2 + r_w^2)^2}{(r_w^2 + r_s^2)^2},
\]

or, for \( r_w = r_0 = \sqrt{2} r_s \), say, \( \gamma = 25\epsilon/144v^2 \), which for \( v = 6 \) and PSF asymmetry \( \epsilon = 0.3 \) (a reasonable value for off-axis points on the CFHT in good seeing) gives a shear of around 1.4%. This is a sizable effect, and should therefore be corrected for.

5. DISCUSSION

We have considered the problem of how to estimate weak gravitational shear from observations that have been degraded by atmospheric and/or instrumental effects. Previous analyses of this problem have made simplifying assumptions that render the results inaccurate. A major result of the paper is the finite-resolution shear operator equation (33), which gives the response of an observed image to a gravitational shear applied before smoothing with the PSF. This result can be used to properly calibrate the effect of any shear estimator, and is valid for an arbitrary PSF, be it turbulence or diffusion limited. We then focused on the application to weighted-moment shear estimators. We have computed the response of individual objects to a shear in § 3.1, and the response of the population of background galaxies with given photometric properties in § 3.2, and from this we have devised an optimal weighting scheme in § 3.3. In the last section we have considered the correction for PSF anisotropy. While there are still some approximations in the present analysis, we feel that they place the techniques of shear measurement on a much firmer footing than before.

I wish to acknowledge helpful discussions with Malcolm Northcott, Francois and Claude Roddier, John Tonry, Ger Luppino, Pat Henry, Ken Chambers, Jeff Kuhn, and Christ Ftaclas.

APPENDIX A

PROPERTIES OF OPTICAL POINT-SPREAD FUNCTIONS

Here we briefly review and derive some properties of telescope PSFs that are used above. For more detailed background, the reader should consult Roddier (1981) and Beckers (1993), and references therein. We will highlight the wavelength or color dependence of the various sources of PSF anisotropy, which may be crucially important for weak-lensing searches for large-scale structure and galaxy-galaxy lensing.

According to elementary diffraction theory (Born & Wolf 1964), the complex electromagnetic field amplitude, \( a(x) \), due to a distant source at position \( x_{\text{phys}} \) on the focal plane (we will suppress polarization subscripts for clarity) is given as an integral
over the input pupil,
\[
a(x_{\text{phys}}) = \int d^2 r A(r) C(r) e^{2\pi i x_{\text{phys}} \cdot r/Ld} ,
\]
(A1)

where \(A(r)\) is the “pupil function” describing the aperture transmission, \(C(r)\) is the complex electric field amplitude of the incoming wave, \(\lambda\) is the wavelength of the radiation, and \(L\) is the focal length. The field amplitude \(C(r)\) will incorporate any random amplitude and phase variations of the incoming wavefronts due to atmospheric turbulence, whereas constant wavefront distortions due to aberrations in the optical elements of the telescope are incorporated as a complex factor in \(A\). Thus, \(a(x)\) is the Fourier transform of \(AC\), evaluated at wavenumber \(k = 2\pi x/L\). The PSF \(g\) is the square of the field amplitude, and in rescaled coordinates \(x = 2\pi x_{\text{phys}}/Ld\) is
\[
g(x) = |a(x_{\text{phys}})|^2 = \int \frac{d^2 z}{(2\pi)^2} e^{-ixz}{\tilde{g}(z)} ,
\]
where the OTF is
\[
\tilde{g}(z) = \int d^2 r C(r) C^*(r + z) A(r) A^*(r + z) .
\]
(A2)

For very short ground-based observations, the atmospheric rippling is frozen and the PSF consists of speckles. For long exposures, we take the time average of the OTF and can replace \(C(r)C^*(r + z)\) by its time average, \(\langle C(r)C^*(r + z)\rangle = \xi(z)\). This is independent of \(r\), so the OTF factorizes into two independent functions,
\[
\tilde{g}(z) = \xi(z) \int d^2 r A(r) A^*(r + z) ,
\]
(A4)

and the same is true for random small-scale amplitude or phase fluctuations introduced by, e.g., random fine-scale mirror roughness.

A1. ATMOSPHERIC TURBULENCE

Ground-based observations on large telescopes are usually limited by atmospheric seeing arising from inhomogeneous random turbulence, and it is a good approximation to ignore the finite size of the entrance aperture and set the factor involving \(A\) in equation (A4) to unity. In the “near field” immediately behind the turbulent layer, the effect on the incoming wave is a pure phase shift, \(C(r) = e^{i\phi(r)}\), where \(\phi = 2\pi d(r)/\lambda\), and \(d(r)\) is the vertical displacement of the wavefront due to turbulence. The displacement \(d\) is nearly independent of wavelength, so \(\phi(r) \propto 1/\lambda\). At greater depths, this phase shift evolves into a combination of amplitude and phase variations, but the two-point function \(\langle C(r)C^*(r + z)\rangle\) remains invariant (Fried 1966; Roddier 1981), and the “natural-seeing” OTF is
\[
\tilde{g}(z) = \langle e^{i(\phi(r) - \phi(r + z))}\rangle .
\]
(A5)

For steady turbulence and long integrations, the central limit theorem guarantees that the phase error \(\psi = \phi(r) - \phi(r + z)\) will have a Gaussian probability distribution \(p(\psi) = (2\pi)^{-1/2} \exp(-\psi^2/2\langle\psi^2\rangle)\), and so the time average of the complex exponential is
\[
\langle e^{i\psi}\rangle = \int d\psi p(\psi)e^{i\psi} = \exp\left(-\frac{\langle\psi^2\rangle}{2}\right) = \exp\left[-\frac{S_\phi(r)}{2}\right] ,
\]
(A6)

where the “phase structure function” is
\[
S_\phi(\Delta r) = \langle (\phi - \phi_0)^2 \rangle .
\]
(A7)

There are strong theoretical (Tatarski 1961) and empirical reasons to believe that on scales much less than some outer scale, the turbulence will have the Kolmogorov \(n = -11/3\) spectrum, for which \(S_\phi \propto r^{5/3}\). This is conventionally written as \(S_\phi(r) \sim 6.88(r/r_0)^{5/3}\), where \(r_0\) is the Fried length (Fried 1966), which is on the order of tens of cm for typical observing conditions (an \(r_0\) of 20 cm gives a FWHM = 0.5 at \(\lambda = 550\) nm). The rms phase difference rises with separation as \(r^{5/6}\) in the “inertial range” delimited at the upper end by the outer scale, set by the width of the mixing layer, which recent estimates (Avila et al. 1997; Martin et al. 1998) find to be around 10–20 m and much larger than \(r_0\). The on-axis OTF computed from stars in deep CFHT imaging agrees quite well with the theoretical expectation. The inertial range is limited at the low end by diffusion, but at scales much smaller than \(r_0\) little error is incurred in ignoring this; in real telescopes, mirror roughness and other effects modify the OTF at small scales. These effects dominate the PSF at very large radii, but are unimportant for weak-lensing observations.

The optical transfer function is \(\tilde{g}(k) = \exp[-S_\phi(kLd/(2\pi)/2)]\), and is real and positive. The PSF is the transform of \(\exp[-6.88(z/r_0)^{5/3}]\), with a width that scales as \(R_{\text{PSF}} \propto \lambda^{-1/5}\), which again is found to apply quite well in practice. This very weak dependence on wavelength is a blessing in weak lensing, since one uses stars to measure the PSF for the galaxies, yet the stars and the galaxies may have different colors. The atmospheric PSF is expected to be isotropic. At large angles, the PSF has a profile \(q \propto x^{-11/3}\), so the unweighted second moment of the PSF is not well defined. For Kolmogorov turbulence, the log of the OTF is just proportional to \(k^{5/3}\), and is well defined for all \(k\).
where depends largely on the altitude of the turbulent layer. The isokinetic patch size (over which stars move coherently) is PSF for large telescopes will be rather modest, at least if the current, rather low, estimates of the outer scale are correct. Fast and where Examples are shown in Figure 6. These plots show that the impact of fast guiding on the atmospheric

for $h$

fast guiding does not cure PSF anisotropies from telescope aberrations. \[ \text{A2. FAST GUIDING} \]

According to the Kolmogorov law, the rms wave-front tilt, averaged over scale $r$, varies as $r^{-1/6}$, which suggests that even for telescopes with reasonably large diameter—Fried length ratio, $D/r_0$, there may be a useful gain in image quality from fast guiding, and experience with HRCAM on the CFHT (McClure et al. 1989) would seem to support this, although part of the dramatic improvement is likely due to the inadequacy of the existing slow guiding system. One technological advance that may have implications for weak lensing is the advent of on-chip fast guiding (Tonry, Burke, & Schechter 1997) with OTCCD chips. With a mosaic camera composed of such devices, it should be possible to obtain partial image compensation over a large angular scale, with one or more guide stars for each “isokinetic” patch.

The theoretical fast-guiding PSF was first explored by Fried (1966), who argued that the OTF should take the form of the natural seeing or uncompensated OTF times an “inverse Gaussian,” $\exp (-zk^2)$, with scale factor $z$ given in terms of $D$ and $r_0$. This is a physically reasonable picture, since it implies that the natural PSF is the convolution of the corrected PSF with a Gaussian to describe the distribution of tilt, but it is only an approximate result. Modified forms of the Fried approximation have been explored by Young (1974) and Jenkins (1998), and the fast-guiding OTF has been simulated by Christou (1991).

It can be shown that in the near-field limit, the exact fast-guiding OTF is given by

$$
\tilde{g}(z) = \int d^2r A(r)A(r+z) \exp \left[-\frac{\langle \psi(r, z)^2 \rangle}{2}\right],
$$

where

$$
\langle \psi(r, z)^2 \rangle = S_{\phi}(z) + z\left[(W_i \otimes S_{\phi})_r - (W_i \otimes S_{\phi})_{r+z}\right]
$$

and where $W_i \equiv \delta_i A^2$. Examples are shown in Figure 6. These plots show that the impact of fast guiding on the atmospheric PSF for large telescopes will be rather modest, at least if the current, rather low, estimates of the outer scale are correct. Fast guiding may, however, yield dramatic improvements for small ($\sim 1$ m diameter) telescopes. How well this would work depends largely on the altitude of the turbulent layer. The isokinetic patch size (over which stars move coherently) is $\sim D/h$, so for $h = 10$ km and $D = 1$ m, say, this is on the order of $20^\circ$; the motion needs to be sampled at a rate $\gtrsim v/D$, reflecting the relatively high wind speed at high altitude, and it may then be difficult to find bright enough guide stars. There are strong indications (Christian & Racine 1985; Tonry et al. 1997; McClure et al. 1991) that centroid motions are coherent over much larger angular scales than this, indicating that much of the image degradation arises from low-altitude turbulence, and this greatly improves the outlook, since one can determine the local motion by averaging a number of stars, and one can afford to sample at a lower rate. A collection of small telescopes equipped with wide-angle OTCCD cameras could be a formidable instrument for weak lensing or other projects requiring high-resolution imaging over wide fields. Fast guiding, while offering important resolution gains, will also present its own challenges, since one expects the PSF to become systematically anisotropic depending on location with respect to the guide stars for the reasons described by McClure et al. (1991). In addition, fast guiding does not cure PSF anisotropies from telescope aberrations.

\[ \text{A3. ATMOSPHERIC DISPERSION} \]

Variation of the refractive index of the atmosphere with wavelength will cause images to be dispersed into spectra, and the consequent elongation of images was estimated by KSB, who pointed out that this could potentially cause problems when

\[ \text{FIG. 6.—Optical transfer functions and corresponding PSFs given by eqs. (A8) and (A9) for a range of telescope aperture diameters: 1.0 m (solid line), 1.5 m (dashed line), 2.2 m (dot-dashed line), 3.6 m (dotted line), and } \infty \text{ (triple-dot-dashed line). A von Karman turbulence spectrum with outer scale of 20 m and Fried length } r_0 = 0.2 \text{ m were assumed, and the telescope was assumed to be operating at a wavelength of } \lambda = 550 \text{ nm.} \]
using PSF’s of stars to correct the shapes of galaxies, since the elongation depends on the spectrum of the object; an object with a linelike spectrum, or one with a sharp-edged absorption band that falls within the bandpass, will of course be less dispersed than a continuum object. Here we will make this more quantitative.

Weak-shear measurements are usually made on images composed from multiple exposures taken at a range of air masses and that are coregistered using centroids of foreground stars. The astrometric solution obtained using stars of a variety of spectra will be a compromise. Assuming, as is usually the case, that one has many more stars at one’s disposal than the number of coefficients of the transformation one is trying to solve for (typically a low-order polynomial), then the solution obtained by minimizing least-squared residuals will be one that correctly maps surface brightness at some wavelength \( \lambda_0 \), determined by the average properties of the foreground stars, but with photons at other wavelengths displaced by an amount proportional to \( \lambda - \lambda_0 \). The sum of a set of \( N \) such images is

\[
f_{\text{tot}}(r, \lambda) = \sum_{i=1}^{N} f(r - \alpha_i(\lambda - \lambda_0), \lambda),
\]

where \( f(r, \lambda) \) is the image seen at zenith. The 2-vector \( \alpha_i \) has \( |\alpha_i| = (\Theta_0/\lambda_0)(\partial \ln \Theta/\partial \ln \lambda) \tan \theta_i \), where \( \theta_i \) is the zenith distance, and \( \hat{\alpha}_i \) is directed toward the horizon. Tables of the refraction angle \( \Theta(\lambda) \) are given by Allen (1973). A star with spectral energy distribution (SED) \( f_s \), and therefore with photon distribution function \( S(\lambda) d\lambda \propto \lambda f_s d\lambda \), gives, in a photon-counting system, a response at zenith \( f(r, \lambda) = \delta(r) S(\lambda) \), and the summed image of such a star is then

\[
f_{\text{tot}}(r) = \int d\lambda f_{\text{tot}}(r, \lambda) = \sum_{i=1}^{N} \int d\lambda \delta(r - \alpha_i(\lambda - \lambda_0)) S(\lambda).
\]

The centroid of a star in the image \( f_{\text{tot}}(r) \) is an average of the centroids that would be obtained from each of the contributing images. Consider the \( I \)th image, and rotate the coordinate system so that \( \alpha_I \) lies along the \( x \)-axis. The centroid for this component is \( \bar{r}_I = (\bar{x}_I, 0) \), with

\[
\bar{x}_I = \frac{\int dx \int dy \int d\lambda \delta(x - \alpha_I(\lambda - \lambda_0)) \delta(y) S(\lambda)}{\int dx \int dy \int d\lambda \delta(x - \alpha_I(\lambda - \lambda_0)) \delta(y) S(\lambda)} = \frac{\int d\lambda \delta(\lambda - \lambda_0) S(\lambda)}{\int d\lambda S(\lambda)} = \alpha_I(\lambda - \lambda_0),
\]

from which it follows that the \( \lambda_0 \) that minimizes \( \langle \bar{x}_I^2 \rangle \), being the average over the stars used for registration of the squared displacement, is \( \lambda_0 = \langle \lambda \rangle = (1/N_{\text{stars}}) \sum_i \bar{x}_I \), i.e., simply the average over the registration stars of their mean wavelength. The centroid for the summed image is

\[
\bar{r}_I = (\bar{x}_I - \lambda_0) \langle \alpha_i \rangle_{1I},
\]

where \( \langle \alpha_i \rangle_{1I} = (1/N) \sum \alpha_i \). Similarly, the central second moment is

\[
p_{1IJ} = \frac{\int d^2r (r_I - \bar{r}_I)(r_J - \bar{r}_J) f_{\text{tot}}(r)}{\int d^2f_{\text{tot}}(r)} = (\lambda - \lambda_0)^2 \langle \alpha_i \alpha_j \rangle_{1I} - (\bar{x}_I - \lambda_0)^2 \langle \alpha_i \rangle_{1I} \langle \alpha_j \rangle_{1I}.
\]

Note that since, for any sensible zenith angle, the width of the spectrum is tiny compared to that of the PSF, this unweighted second moment fully characterizes the effect of dispersion.

Consider, for illustration, the case of an equatorial field observed with a telescope near the equator, and for a range of zenith angles \( |z| < z_{\text{max}} \). In this case, \( \sum \alpha_i = 0 \), so the second term in equation (A14) vanishes (this is not a particularly unusual special case; for a number of observations spread over a range of zenith angles, one would generally expect the second term here to become relatively small). When we solve for the PSF from the shapes of foreground stars of various types, we are effectively averaging over a mix of SEDs appropriate for a low-redshift spiral galaxy. The average PSF moment obtained from the stars can then be written, for the equatorial case,

\[
\langle p_{xx} \rangle = \langle \bar{x}_I^2 \rangle - \langle \bar{x}_I \rangle^2 \langle x^2 \rangle_{1I},
\]

whereby \( \langle x^2 \rangle_{1I} = (\Theta_0/\lambda_0)^2(\partial \ln \Theta/\partial \ln \lambda)^2(1/N) \sum_i \tan^2 z_i \) and the SED dependent coefficient of \( \langle x^2 \rangle_{1I} \) is the mean of the dispersions of the stars plus the dispersion of the means of the stars. For the equatorial case, this is the same as the second moment for a single pointlike object with an SED like the average SED of the registration stars. This is not strictly true in general, but the difference is probably a minor one, and it then follows that if the faint galaxies have SEDs such that their \( (\lambda - \lambda_0)^2 \) differs systematically from that for a low-redshift spiral, then the PSF correction will be systematically in error. Figure 7 quantifies the importance of this effect using redshifted galaxy SEDs of various types from Coleman, Wu, & Weedman (1980). The left panel shows the displacement of the centroid of galaxies of the various indicated types as their spectra are redshifted, and the right panel shows the variation of the second moment of the PSF. We find that the \( I \)-band second moment is very stable, with peak fluctuations of \( \delta p_{xx} \sim 300 \text{ mas}^2 \) for these parameters, and with little systematic difference from a low spiral redshift galaxy if we integrate over a range of redshifts. For illustration, if we take the systematic change in polarization to be, say, \( \delta p_{xx} \sim 100 \text{ mas}^2 \) and a Gaussian-profile galaxy, this corresponds to a shear of \( \gamma \approx 1.38 \delta p_{xx} / \text{FWHM}^2 \approx 6 \times 10^{-4}(0.5/\text{FWHM})^2 \). In the \( V \) band the polarization fluctuations are larger by a factor of 2–3, but even then, and even for marginally resolved objects in excellent seeing, the effect is at or below the sensitivity for current and
Fig. 7.—Atmospheric dispersion. Left: Angular displacement due to changes in the SED for galaxies at a range of redshifts. A zenith angle of 45° was assumed; top and bottom panels show the deflection for the $I$-band and $V$-band system response functions, respectively, for the CFH 12K camera. More concretely, the quantity plotted is the mean wavelength where is the SED, $R(\lambda)$ is the system response, $\lambda_0$ was taken to be 820 nm, and $\lambda - \lambda_0$ was converted to angle as described in the text, and assuming a 4000 m observatory altitude. The plots show that the shifts in the centroids due to SED variation are on the order of 20 mas in the $I$-band, but somewhat larger in $V$. The elliptical/S0 SED shows a somewhat greater excursion at high redshift, but this is caused by the SED redshifting right out of the filter band, so one would not expect to find many such objects in flux-limited samples. Right: plots showing how the width of the spectrum, and hence the PSF polarization, varies with redshift and galaxy type. Here the quantity plotted is again converted to angle, here assuming $\tan^2 z_T \approx 1$. 

A4. ABERRATIONS

Aberrations of the optical elements of the telescope can be a significant contribution to the anisotropy of the PSF. These can be analyzed in the same manner as the wavefront deformation due to the atmosphere, but with a couple of distinctive features: first, for low-order “classical” aberrations, where the phase error varies smoothly, and for ground-based observing conditions, if an aberration is an important factor, then its contribution to the OTF will be nearly achromatic. This is because if there is a smooth variation of the wavefront error (rather than a Gaussian random field with power at all scales, as in atmospheric turbulence) amounting to $N \approx 1$ wavelengths, then the PSF will be very well approximated by its geometric optics limit, with a shape defined by the pattern of caustics (although for a narrowband filter, the PSF would actually be found on close examination to be composed of a set of speckle-sized patches concentrated along the lines where the classical caustics form; Berry & Upstill 1980). The wavefront deformation can be measured directly from out-of-focus images (Roddier & Roddier 1993), so this contribution to the PSF can be directly predicted.
Truly diffraction limited seeing arises when the rms phase error across the aperture (due to the atmosphere and/or aberration of the optical elements of the telescope) is much less than unity. In this case, the optical transfer function, \( \tilde{g}(k) \), is to a good approximation, just the autocorrelation of the aperture, \( A \otimes A \), at lag \( \Delta r = kD\lambda/2\pi \), and must therefore vanish for spatial frequencies \( k > 2\pi/(f\lambda) \), where \( f \) is the ratio of the aperture diameter to the focal length. The log of the OTF becomes ill defined as one approaches the diffraction limit. From equation (A4), we see that this cutoff is also present in the case of turbulence-dominated seeing, but it occurs at a high frequency, where the optical transfer function has already become exponentially small due to atmospheric effects, and has little impact.

In the absence of aberrations, the OTF for diffraction-limited seeing is real and nonnegative. The OTF is symmetric under rotations of 180°, and so any quadrupole anisotropy of the PSF can be annulled simply by reconvolving one’s image with a 90° rotated PSF. For diffraction-limited seeing, the size of the PSF scales as the inverse of the wavelength, a fact that can be incorporated into empirical or theoretical (Krist 1995) modeling of the PSF.

In the HST WFPC2, figure errors are not negligible. The imaginary part of the OTF is excited to the degree that reconvolution with a 90° PSF still leaves nonnegligible PSF anisotropy. The phase errors are not large (the telescope is nearly diffraction limited), which means that the wavelength dependence could be quite complicated. The systematic error arising from differences between faint galaxy and foreground star SEDs can be estimated much as we did for atmospheric dispersion.

So far, we have considered the continuous distribution of intensity on the focal plane, \( f_p(r) \). In real detectors, we sample the image with a grid of pixels. The response of a pixel is not uniform and has been directly measured for front-illuminated EEV images to a small fraction of a pixel (say, \( \pm 0.05 \) pixels), and the effect of inaccuracy at this level will have a negligible effect on atmospheric, telescope transfer functions with the Fourier transform of this pattern. In practice, one can typically register sky. Each image gives a two-dimensional grid of samples of \( I \), and a piecewise continuous function can be constructed by shifting the grid of \( \delta \)-functions into an absolute astrometric coordinate system and convolving with some interpolation function, \( \mathcal{P}_\text{interp} \). One can incorporate the effect of guiding errors on a single exposure as a convolution with the pixel function. If we average a set of \( N \) such images, the result is

\[
F(r) = \frac{1}{N} \sum \left[ (f_p \otimes p) \cdot c_A \right] \otimes \mathcal{P}_\text{interp},
\]

where \( c_A(r) \) is a two-dimensional comb function with spatial offset (in units of the pixel spacing) \( \Delta \):

\[
c_A(r) = \sum_{i_s,i_p=\pm\infty} \delta(r - (i + \Delta)) .
\]

The form of \( \mathcal{P}_\text{interp} \) depends on the type of interpolation used. For “nearest pixel” interpolation, \( \mathcal{P}_\text{interp} \) is just a uniform box of side \( d \), but if one linearly interpolates between the pixel samples, for example, then \( \mathcal{P}_\text{interp} \) will be a more extended, but again readily computable, function.

The transform of \( F(r) \) is

\[
\tilde{F}(k) = \mathcal{P}_\text{interp} \sum_{m_x,m_y=\pm\infty} (\tilde{f}_p \tilde{p})_{k-2\pi m/d} \frac{1}{N} \sum \exp(2\pi im \cdot \Delta) .
\]

The \( m_x = m_y = 0 \) term in the first sum here is just the ideal image \( f_p \) convolved with \( p \) and with \( \mathcal{P}_\text{interp} \), while the higher order terms represent aliasing. The transform \( \tilde{F}(k) \) is \( \mathcal{P}_\text{interp} \) times the superposition of a grid of images of \( \tilde{f}_p \), with spacing \( 2\pi/d \). Since \( \tilde{p} \) is a fairly compact function, the dominant aliasing comes from the low-order images, \( m = \pm 1 \). Aliasing is most severe for a single exposure, since the low-order aliased images contribute with unit weight. If we average \( N \) randomly shifted images, then the strength of the \( m \neq 0 \) terms is reduced by a factor of \( \sim N^{-1/2} \), aliasing is greatly reduced, and to a good approximation the field \( F(r) \) is simply the convolution of \( f_p \) with \( p \otimes \mathcal{P}_\text{interp} \). With systematically staggered images, as is possible with HST and potentially with fast on-chip guiding, one can do even better; with a uniform \( M \times M \) grid of offsets covering the unit pixel, the nearest, and therefore most problematic, modes, \( m_x, m_y = \pm 1 \), are then zero, and the modes remain small until we get to a multiple of \( 2M \) times the Nyquist frequency. This assumes that the transformation from detector to sky coordinates is determined and applied accurately. If we make finite errors \( \delta \Delta \) in registration, the resulting image will be the convolution of the ideal PSF with a highly compact cluster of \( \delta \)-functions, and the optical transfer function, \( \tilde{g} \), will be the product of the atmospheric, telescope transfer functions with the Fourier transform of this pattern. In practice, one can typically register images to a small fraction of a pixel (say, \( \leq 0.05 \) pixels), and the effect of inaccuracy at this level will have a negligible effect on the final PSF.

Noise in the images, assumed to be incoherent Poisson noise in the source images, can be analyzed in a similar manner, and we find that the two-point function of the noise is just the convolution of \( \mathcal{P}_\text{interp} \) with itself, and the two-point function of the noise in recircularized images can be obtained by convolving the raw noise ACF with \( g^2 \) twice.
APPENDIX B

SIMULATED DATA

To test the procedures described here, we have generated simulated mock data and then analyzed these. The simulations were made to match as closely as possible observations of ~3 hr integration on the CFHT with 0′.6 seeing.

We first generated a set of 200 mock catalog of galaxies, each corresponding to a patch of sky of size 2′56 on a side. Galaxies were drawn from a Schechter-style luminosity function, laid down in a Poissonian manner in an Einstein–de Sitter cosmology. Images with pixel scale 0′.075 were then generated by realizing the galaxies as exponential disks with random orientations and scale lengths corresponding to fixed rest frame central surface brightness. The galaxies were modeled as optically thick, since the optically thin model looks unrealistic, having too many very bright edge-on systems compared to real images. A number of pointlike stars were added to the images, which were then sheared with γ = 0.1, convolved with a Kolmogorov turbulence PSF with 0′.6 FWHM, and then rebinned to 0′.15 pixel scale. When the real data are analyzed, they are interpolated from the original 0′.2 pixel scale to the final 0′.15 scale with bilinear interpolation. This results in a further convolution of the signal and the noise in the real images, but with slightly different smoothing kernels. These kernels were computed by modeling the image shifts as a uniform distribution within the final pixel size: the noise-free mock images were convolved with the appropriate kernel, and then Gaussian white-noise images were generated to model the sky noise, and were convolved with the appropriate kernel and added to the images. A sample image is shown in Figure 8, and the corresponding size-magnitude diagram is shown in Figure 9, from which it is apparent that the simulated objects have properties very similar to those detected in the real data.

These data were analyzed exactly like the real data. That is, the objects were detected as peaks of a smoothed image. The stars were extracted and their shapes fitted in the manner described to obtain the PSF. A smoothed image \( f_s = g \odot f_o \) was...
generated, and from this the polarization $q_s$ was computed using equation (45), and the polarizability for each object was computed using equation (48).

REFERENCES

Allen, C. 1973, Astrophysical Quantities (London: Althone)
Avila, R., Ziad, A., Borgnino, J., Martin, F., Agabi, A., & Tokovinin, A. 1997, J. Opt. Soc. Am., 14, 3070
Bar-Kana, R. 1996, ApJ, 468, 17
Beckers, J. 1993, ARA&A, 31, 13
Berry, M. V., & Upstill, C. 1980, Prog. Opt., 18, 257
Bertin, E., & Arnouts, S. 1996, A&AS, 117, 393
Bonnet, H., & Mellier, Y. 1995, A&A, 303, 331
Born, M., & Wolf, E. 1964, Principles of Optics (Oxford: Pergamon)
Christian, C., & Racine, R. 1985, PASP, 97, 1215
Christou, J. C. 1991, PASP, 103, 1040
Coleman, G. D., Wu, C.-C., & Weedman, D. W. 1980, ApJS, 43, 393
Fischer, P., & Tyson, J. A. 1997, AJ, 114, 14
Fried, D. L. 1966, J. Opt. Soc. Am., 56, 1372
Gunn, J. 1967, ApJ, 147, 61
Hoekstra, H., Franx, M., Kuijken, K., & Squires, G. 1998, ApJ, 504, 636
Jarvis, J. F., & Tyson, J. A. 1981, AJ, 86, 476
Jenkins, C. R. 1998, MNRAS, 294, 69
Jorden, P., Deltorn, J., & Oates, P. 1993, Gemini, 41, 1
———. 1994, Proc. SPIE, 2198, 836
Kaiser, N. 1998, ApJ, 498, 26
Kaiser, N., Squires, G., & Broadhurst, T. 1995, ApJ, 449, 460 (KSB)
Kaiser, N., Wilson, G., Luppino, G., Kofman, L., Gioia, I., Metzger, M., & Dahle, H. 1998, ApJ, submitted (preprint astro-ph/9809268)
Kochanek, C. 1990, MNRAS, 247, 135
Krist, J. 1995, ASP Conf. Ser. 77, Astronomical Data Analysis Software and Systems IV, ed. R. A. Shaw, H. E. Payne, & J. J. E. Haynes (San Francisco: ASP), 349
Luppino, G., & Kaiser, N. 1997, ApJ, 475, 20
Martin, F., Tokovinin, A., Ziad, A., Conner, R., Borgnino, J., Avila, R., Agabi, A., & Sarazin, M. 1998, A&A, 336, L49
McCune, R. D., Fletcher, J. M., Arnaud, J., Nieto, J. L., & Racine, R. 1991, PASP, 103, 570
McCune, R. D. et al. 1989, PASP, 101, 1156
Roddier, C., & Roddier, F. 1993, J. Opt. Soc. Am., 10, 2277
Roddier, F. 1981, Prog. Opt., 19, 281
Schneider, P., Ehlers, J., & Falco, E. E. 1992, Gravitational Lenses (New York: Springer)
Tatarski, V. I. 1961, Wave Propagation in a Turbulent Medium (New York: McGraw-Hill)
Tonry, J., Burke, B. E., & Schechter, P. L. 1997, PASP, 109, 1154
Tyson, J. A., Valdes, F., Jarvis, J. F., & Mills, A. P., Jr. 1984, ApJ, 281, L59
Valdes, F., Jarvis, J. F., & Tyson, J. A. 1983, ApJ, 271, 431
Wilson, G., Cole, S., & Frenk, C. S. 1996, MNRAS, 280, 199
Young, A. T. 1974, ApJ, 189, 587