The Estimation of Object's Ellipticities

Robert Lupton

May 21, 1996

1 Approaches to Measuring Image Shapes

The estimation of the axis ratio and position angle of an object detected on a CCD image is not straightforward. The approaches used include:

1. The shape of some specified isophote.
2. Unweighted second moments of x and y within a specified isophote.
3. Second moments of x and y, weighted by the intensity, without any boundary.
4. Second moments of x and y, weighted by the intensity, within a specified isophote.
5. Second moments of $x/f(r)$ and $y/f(r)$, weighted by the intensity, without any boundary.
6. Second moments of $x/f(r)$ and $y/f(r)$, weighted by the intensity, within some boundary.
7. Fit a model to the 2-dimensional data. (This is not yet discussed here).

Note that the moments used in method 3 diverge for profiles that don't fall off faster than r^{-4}; we shall therefore exclude it from further consideration.

Methods 1 and 2 are merely different ways of measuring the same thing. If the isophote level is chosen to be close to sky, the shape of the isophote is rather unstable; depending on exactly what level is chosen (and e.g. how the sky is estimated), large portions of low-surface-brightness features may be in- or excluded essentially at random.

Another consequence of the divergent nature of image moments is that measures such as method 4 are dominated by the pixels very close to the chosen isophote; in fact all of the isophote-based measures are basically measuring the same thing.

In the SDSS's photometric pipeline, we have chosen to use method 6 with $f(r) = r$. Although in a noise-free world we would prefer method 5, it is untenable; although the portions of the image outside the object contribute no signal to the moments, they do contribute noise. For the SDSS we choose a circular boundary containing almost all of the light, and expect that the resulting truncation of the object's profile will have little effect on the measured shape.

2 Properties of inverse-radius weighted moments

Let use define

$$\begin{aligned} M_{xx} &\equiv \langle x^2/r^2 \rangle \\ M_{xy} &\equiv \langle xy/r^2 \rangle \\ M_{yy} &\equiv \langle y^2/r^2 \rangle, \end{aligned}$$

where the moments are weighted by the counts.

Firstly, note that these moments exist if the energy in an object is finite, and that their largest contribution comes from the part of the object that contributes the most light; both of these are desirable properties. Note also that $M_{xx} + M_{yy} \equiv 1$.

If the object's isophotes are self-similar ellipses, we may readily calculate these moments:

$$M_{xx} \equiv \frac{I_{20}}{I_{00}}$$

where

$$I_{\zeta\eta} \equiv \int_{-\infty}^{\infty} (x/r)^{\zeta} (y/r)^{\eta} \mu(x,y)\, dx\, dy$$

Consider for a moment the case where the major and minor axes of the ellipse are aligned with the x and y axes. If the surface density is $\mu(x, y)$, we can change variables to a, θ defined by

$$x = a\cos\theta$$
$$y = b\sin\theta \equiv a\alpha\sin\theta$$

with Jacobian

$$\tfrac{\partial(x,y)}{\partial(a,\theta)} = a\alpha$$

and find that

$$\begin{aligned} I_{\zeta\eta} &= \int_0^\infty \mu(a)\, 2\pi a\, da \cdot \frac{1}{2\pi}\int_0^{2\pi} \frac{\cos^\zeta\theta\sin^\eta\theta}{(\cos^2\theta + \alpha^2\sin^2\theta)^{(\zeta+\eta)/2}} d\theta \\ &\equiv F(\mu)\cdot J_{\zeta\eta}(\alpha). \end{aligned}$$

The $J_{\zeta\eta}$ are readily evaluated via a $t = \tan\theta$ substitution, leading to

$$J_{00} = \alpha$$

$$J_{10} = 0 \quad J_{01} = 0$$

$$J_{20} = \frac{\alpha}{1+\alpha} \quad J_{11} = 0 \quad J_{02} = \frac{\alpha^2}{1+\alpha}$$

$$J_{40} = \frac{\alpha(\alpha+2)}{2(1+\alpha)^2} \quad J_{22} = \frac{\alpha^2}{2(1+\alpha)^2} \quad J_{04} = \frac{\alpha^2(1+2\alpha)}{2(1+\alpha)^2}$$

We can now find M_{xx} and M_{yy}:

$$M_{xx} = \frac{I_{20}}{I_{00}} = \frac{J_{20}}{J_{00}} = \frac{1}{1+\alpha}$$

and

$$M_{yy} = \frac{I_{02}}{I_{00}} = \frac{J_{02}}{J_{00}} = \frac{\alpha}{1+\alpha}$$

The form of the radial profile does not enter into these results; this may be less surprising once you remember that for round objects $M_{xx} = M_{yy} = 1/2$.

We see that the axis ratio α is given by

$$\alpha = \frac{M_{yy}}{M_{xx}}; \tag{1}$$

note that there are no square roots in this result.

If we consider a different set of axes, given by

$$\begin{aligned} x' &= x\cos\phi - y\sin\phi \\ y' &= x\sin\phi + y\cos\phi \end{aligned}$$

we find that

$$\begin{aligned} M_{x'x'} &= M_{xx}\cos^2\phi + M_{yy}\sin^2\phi \\ &= \frac{1}{1+\alpha}\left(\cos^2\phi + \alpha\sin^2\phi\right) \\ M_{x'y'} &= (M_{xx} - M_{yy})\sin 2\phi/2 \\ &= \frac{1-\alpha}{1+\alpha}\sin 2\phi \\ M_{y'y'} &= M_{xx}\sin^2\phi + M_{yy}\cos^2\phi \\ &= \frac{1}{1+\alpha}\left(\sin^2\phi + \alpha\cos^2\phi\right), \end{aligned}$$

A little algebra then shows that

$$\frac{a-b}{a+b}\cos 2\phi \equiv Q = M_{x'x'} - M_{y'y'} \equiv 2M_{x'x'} - 1 \tag{2}$$

$$\frac{a-b}{a+b}\sin 2\phi \equiv U = 2M_{x'y'} \tag{3}$$

from which we may easily derive the axis ratio and position angle. The quantities Q and U are named by analogy to the Stoke's parameters used to describe polarization; we may also define $P \equiv (a-b)/(a+b)$.

3 Statistical Properties of Weighted Moments

We have defined

$$I_{\zeta\eta} \equiv \int_{-\infty}^{\infty} (x/r)^\zeta (y/r)^\eta \mu(\mathbf{x})\, d\mathbf{x}.$$

Writing $\mu(\mathbf{x}) = \mu_0(\mathbf{x}) + \epsilon(\mathbf{x})$ with $\langle\epsilon\rangle = 0$ and $\langle\epsilon(\mathbf{x})\epsilon(\mathbf{y})\rangle = E(\mathbf{x})\delta(\mathbf{x}-\mathbf{y})$, and reverting to coordinates aligned with the object's principal axes, we can find

I's mean and variance; as before we'll assume that μ is constant on ellipses:

$$\begin{aligned}\langle I_{\zeta\eta}\rangle &= \int_{-\infty}^{\infty}(x/r)^{\zeta}(y/r)^{\eta}\mu_0(\mathbf{x})\,d\mathbf{x} \\ &= F(\mu_0)J_{\zeta\eta}(\alpha).\end{aligned}$$

The covariances are a little more complicated, as both the object and the sky contribute noise. Let us assume that the radial integrals are cut off at a radius R, assumed larger than the object under study and that the sky surface brightness is S so $E(\mathbf{x}) = \mu_0(\mathbf{x}) + S$, then we have

$$\begin{aligned}\langle (I_{\zeta\eta} - \langle I_{\zeta\eta}\rangle)(I_{\nu\xi} - \langle I_{\nu\xi}\rangle)\rangle &= \int_{-\infty}^{\infty}(x/r)^{\zeta+\nu}(y/r)^{\eta+\xi}E(\mathbf{x})\,d\mathbf{x} \\ &= F(\mu_0)J_{\zeta+\nu,\eta+\xi}(\alpha) + \pi R^2 S J_{\zeta+\nu,\eta+\xi}(0).\end{aligned}$$

4 Statistics of Derived Quantities

The moments $M_{x'x'}$, $M_{x'y'}$, and $M_{y'y'}$ are the ratios of Gaussian (actually, Poisson) variables, and their statistical properties are discussed in section. A summary would be that in general such ratios have poor statistical properties, but if the denominator (the total flux) has a reasonable signal-to-noise ratio τ they are reasonably well behaved.

Q and U are of course modified by rotations of our coordinate system (or, equivalently, of the object), but it is clear that their biases and variances are *not* altered by such coordinate transformations. This means that we can perform the statistical analysis in coordinates aligned with the major and minor axis of the galaxy.

We have chosen to calculate the $I_{\zeta\eta}$ within a circular aperture, whereas the object is flattened, and the photon noise from the sky and the object therefore enter in rather different ways. This means that the results for the variances depend on whether the sky or the object noise dominates; fortunately the differences are minor. The signal-to-noise parameter τ is given by

$$\tau = \frac{\alpha F(\mu_0)}{\sqrt{\alpha F(\mu_0) + \pi R^2 S}}.$$

In the limit where sky noise dominates, the results from the next section reveal that:

$$\langle Q\rangle = \frac{1-\alpha}{1+\alpha}\left(1 + \tau^{-2} + 3\tau^{-4} + O(\tau^{-6})\right)$$

$$
\begin{aligned}
&= P\left(1+\tau^{-2}+3\tau^{-4}+O(\tau^{-6})\right)\\
\langle U\rangle &= 0\\
\mathrm{Var(Q)} &= \frac{3\alpha^2-2\alpha+3}{2(1+\alpha)^2}\tau^{-2}+\frac{19\alpha^2-26\alpha+19}{2(1+\alpha)^2}\tau^{-4}+O(\tau^{-6})\\
&= \frac{1+2P^2}{2}\tau^{-2}+\frac{3+16P^2}{2}\tau^{-4}+O(\tau^{-6})\\
\mathrm{Var(U)} &= \frac{1}{2}\tau^{-2}+\frac{3}{2}\tau^{-4}+O(\tau^{-6})\\
\langle(Q-\langle Q\rangle)(U-\langle U\rangle)\rangle &= 0.
\end{aligned}
$$

In the less interesting regime where the object noise is more important the means and covariance are unaltered and variances become

$$
\begin{aligned}
\mathrm{Var(Q)} &= 2\frac{\alpha^2-\alpha+1}{(1+\alpha)^2}\tau^{-2}+\frac{11\alpha^2-16\alpha+11}{(1+\alpha)^2}\tau^{-4}+O(\tau^{-6})\\
&= \frac{1+3P^2}{2}\tau^{-2}+\frac{3+19P^2}{2}\tau^{-4}+O(\tau^{-6})\\
\mathrm{Var(U)} &= 2\frac{\alpha}{(1+\alpha)^2}\left(1+\tau^{-2}+3\tau^{-4}+O(\tau^{-6})\right)\\
&= \frac{1-P^2}{2}\left(\tau^{-2}+3\tau^{-4}+O(\tau^{-6})\right)
\end{aligned}
$$

Note that, although Q and U are uncorrelated they are not independent, as they are not Gaussian variables.

Rather than Q and U directly, we are often interested in an object's ellipticity and the position angle of its major axis. These are given respectively by

$$
\frac{a-b}{a+b} \equiv P=\sqrt{Q^2+U^2}
$$

and

$$
\tan(2\phi) = \frac{U}{Q};
$$

let us investigate their statistical properties.

Whenever the value of Q is not several standard deviations away from 0, the distribution of ϕ has a significant Cauchy component, so it's not possible to say whether it is unbiased in the usual sense of the term (as it has no expectation value). Its median, on the other hand, is 0 for all values of μ

and σ, so we may reasonably call it unbiased. We can estimate the median's uncertainty from the interquartile range, and the results are given in a table in section 5.

We can easily find the moments of P^2 using the results developed here and in the following section; if the sky is the main source of noise,

$$\begin{aligned}\langle P\rangle &= \frac{(1-\alpha)^2}{(1+\alpha)^2}+4\frac{(\alpha^2-\alpha+1)}{(1+\alpha)^2}\tau^{-2}+O(\tau^{-4})\\ &= P^2+(1+3P^2)\tau^{-2}+O(\tau^{-4})\\ \mathrm{Var}(P) &= 2\frac{(3\alpha^2-2\alpha+3)(1-\alpha)^2}{(1+\alpha)^4}\tau^{-2}+O(\tau^{-4})\\ &= 2P^2(1+2P^2)\tau^{-2}+O(\tau^{-4});\end{aligned}$$

if the noise from the object dominates, the variance becomes

$$\begin{aligned}\mathrm{Var}(P) &= 8\frac{(\alpha^2-\alpha+1)(1-\alpha)^2}{(1+\alpha)^4}\tau^{-2}+O(\tau^{-4})\\ &= 2P^2(1+3P^2)\tau^{-2}+O(\tau^{-4});\end{aligned}$$

The variance is not strongly affected by the dominant noise source.[1]

Although the differences are only significant at small signal-to-noise ratios, $\langle Q\rangle^2+\langle U\rangle^2$ is considerably less biased than $\langle Q^2+U^2\rangle$, implying that we should first take the expectation values of Q and U, and only then square and add. This is analogous to the question of determining the square of the mean of a set of data: $\langle\bar{x}^2\rangle=\mu^2+\sigma^2/n$, while $\langle x^2\rangle=\mu^2+\sigma^2$. If we have only one measurement of an object the distinction is moot.

5 Statistics of the Ratios of Gaussian Variates

Let us initially consider ratios of uncorrelated Gaussian variables. If x_1 is an $N(0,1)$ variable, and x_2 is $N(\mu,\sigma^2)$, then $y\equiv x_1/x_2$ has exact p.d.f.

$$f(y) = \frac{\sigma}{(1+y^2\sigma^2)\,\pi}\,e^{-\frac{\mu^2}{2\sigma^2}}+\frac{\mu}{\sqrt{2\pi}}\frac{1}{(1+y^2\sigma^2)^{3/2}}\,e^{-\frac{\mu^2y^2}{2(1+y^2\sigma^2)}}\mathrm{erf}\left(\frac{\mu}{\sigma\sqrt{2(1+y^2\sigma^2)}}\right).$$

[1] One might have expected that P's variance would be given by $\mathrm{Var}Q+\mathrm{Var}U=P^2(1-P)(1+3P^2)/(1+P)\tau^{-2}+O(\tau^{-4})$, but this is incorrect as U and Q, although uncorrelated, are not independent.

τ	SIQR	σ'	σ''	τ	SIQR	σ'	σ''
0	1.00000	1.48260	0.00000	3	0.23064	1.08131	1.02582
0.125	0.99223	1.48253	0.18389	4	0.17109	1.04585	1.01462
0.25	0.96954	1.48169	0.35936	5	0.13618	1.02946	1.00947
0.375	0.93366	1.47837	0.51909	6	0.11314	1.02029	1.00641
0.5	0.88721	1.47064	0.65769	7	0.09681	1.01492	1.00472
0.75	0.77563	1.43744	0.86246	8	0.08462	1.01146	1.00365
1	0.65999	1.38382	0.97851	9	0.07516	1.00912	1.00295
1.5	0.47271	1.26346	1.05126	10	0.06762	1.00747	1.00247
2	0.35375	1.17275	1.04894	50	0.01349	1.00039	1.00019
2.5	0.27952	1.11585	1.03604	∞	0	1.00000	1.00000

Table 1: The width of the distribution of the ratio x_1/x_2 of two independent Gaussians, $N(0,1)$ and $N(\mu,\sigma^2)$ respectively. The signal-to-noise in the denominator is $\tau \equiv \mu/\sigma$. SIQR is the semi-interquartile range in units of σ. The columns σ' and σ'' are the SIQR divided by 0.6744897502 and multiplied by $\sqrt{1+\tau^2}$ and τ respectively. The numerical factor is that required to convert the SIQR to the standard deviation of a Gaussian.

making the substitution $u = \mu y/(1+y^2\sigma^2)^{1/2}$, and writing $\tau = \mu/\sigma$ (the signal-to-noise-ratio) we have

$$f(u) = \frac{1}{\pi\sqrt{\tau^2-u^2}}e^{-\tau^2/2} + \frac{1}{\sqrt{2\pi}}e^{-u^2/2}\mathrm{erf}\left(\sqrt{\frac{\tau^2-u^2}{2}}\right)$$

where u ranges from $-\tau\cdots\tau$. The first term in this expression is a Cauchy distribution, which possesses no moments, not even a mean, although its median is well defined. The accompanying table gives the semi-interquartile range for a number of values of τ. Unsurprisingly, when μ is many σ away from 0 the distribution's spread becomes that expected for an $N(0,1)$ Gaussian divided by a constant μ.

Because of the factor $\exp(-u^2/2)$ in the last term, we need only consider the error function when u is small; If τ is large we may use an asymptotic expansion, giving

$$f(u) \approx \frac{1}{\sqrt{2\pi}}e^{-u^2/2} + \frac{1}{\pi\sqrt{\tau^2-u^2}}e^{-\tau^2/2}\left(\frac{1}{\tau^2-u^2} - \frac{3}{(\tau^2-u^2)^2}\right)$$

the first term shows that u is an $N(0,1)$ variate, although y itself is not.

For calculation of the properties of the moments M_{uv} we may neglect the Cauchy nature of this distribution, as it is suppressed by a factor $\exp(-\tau^2/2)$ where τ is the signal-to-noise ratio of the object's detection; when we come to consider position angles we shall not be so fortunate.

As an alternative to integrating over the exact p.d.f., it is reasonably easily shown that if u has characteristic function $\phi(t)$, then the moment $\langle 1/u^s \rangle$ is given by

$$\langle 1/u^s \rangle = \frac{1}{i^s(s-1)!} \int_0^\infty t^{s-1} \phi(t) dt$$

the path of integration may in fact be taken to infinity anywhere in the upper half of the complex plane.

We can use this result to compute the inverse-moments of an $N(\mu, \mu^2/\tau^2)$ Gaussian (i.e. $\mu/\sigma = \tau$). Neglecting terms exponentially small in τ^{-2}, and using an asymptotic expansion for $\mathrm{erf}(x)$, we find that:

$$\langle x^{-1} \rangle \approx \frac{1}{\mu} \left(1 + \tau^{-2} + 3\tau^{-4} + 15\tau^{-6} \right) \tag{4}$$

$$\langle x^{-2} \rangle \approx \frac{1}{\mu^2} \left(1 + 3\tau^{-2} + 15\tau^{-4} + 105\tau^{-6} \right) \tag{5}$$

$$\langle x^{-3} \rangle \approx \frac{1}{\mu^3} \left(1 + 6\tau^{-2} + 45\tau^{-4} + 420\tau^{-6} \right) \tag{6}$$

$$\langle x^{-4} \rangle \approx \frac{1}{\mu^4} \left(1 + 10\tau^{-2} + 105\tau^{-4} + 1260\tau^{-6} \right) \tag{7}$$

(parenthetically, I note that these results can also be obtained by expanding $x^{-n} = \mu^{-n}(1 + z/\tau)^{-n}$, where z is an $N(0,1)$ variable; for example $\mu \langle x^{-1} \rangle = \langle 1 \rangle + \langle z \rangle \tau - + \langle z^2 \rangle \tau^{-2} + \langle z^3 \rangle \tau^{-3} + \langle z^4 \rangle \tau^{-4} \cdots \approx 1 + \tau^{-2} + 3\tau^{-4}$; the expansion is only valid for $\tau > 1$, and the series is asymptotic rather than convergent as the coefficient of τ^{-n} is $(n-1)(n-3)\cdots 1$).

In order to lift the restriction that x_1 and x_2 are independent, one way to proceed would be to generalise the result on inverse moments to two variates u_1 and u_2 with a joint characteristic function of $\phi_{12}(t_1, t_2)$; the moment $\langle u_1^r / u_2^s \rangle$ is given by

$$\langle u_1^r / u_2^s \rangle = \frac{1}{i^s(s-1)!} \int_0^\infty t_2^{s-1} \left. \frac{\partial^r \phi_{12}(t_1, t_2)}{\partial (it_1)^r} \right|_{t1=0} dt_2$$

and the resulting algebra is unpleasant. In the case of two correlated Gaussian variables with Pearson coefficient $r_{x_1,x_2} = \rho$, we have an alternative:

because any combination of bivariate Gaussians is itself bivariate Gaussian, we can write

$$x_1 = A + Bx_2 + Cz$$

where z is again an $N(0,1)$ Gaussian independent of x_2, and

$$\begin{aligned} A &= \mu_1 - \mu_2 \rho \sigma_1/\sigma_2 \\ B &= \rho\sigma_1/\sigma_2 \end{aligned}$$

and

$$C = \sigma_1(1-\rho^2)^{1/2}.$$

For example,

$$\begin{aligned} \langle x_1/x_2 \rangle &\approx \frac{\mu_1}{\mu_2} + \left(\frac{\mu_1}{\mu_2} - \rho\frac{\sigma_1}{\sigma_2}\right)\tau^{-2}\left(1+3\tau^{-2}\right) + O(\tau^{-6}) \\ \mathrm{Var}(x_1/x_2) &\approx \left(\frac{\sigma_1^2}{\sigma_2^2} - 2\rho\frac{\mu_1\sigma_1}{\mu_2\sigma_2} + \frac{\mu_1^2}{\mu_2^2}\right)\tau^{-2} + \\ &\qquad \left((3+5\rho^2)\frac{\sigma_1^2}{\sigma_2^2} - 16\rho\frac{\mu_1\sigma_1}{\mu_2\sigma_2} + 8\frac{\mu_1^2}{\mu_2^2}\right)\tau^{-4} + O(\tau^{-6}) \end{aligned}$$

where τ is the signal-to-noise of the denominator, μ_2/σ_2.

It is perhaps a little surprising that finite answers can be obtained from a distribution which has a Cauchy term, and thus infinite moments. To investigate this, consider the mean of an inverse Poisson distribution, whose characteristic function is $\exp(\mu(e^{it}-1))$:

$$\langle 1/x \rangle = \frac{1}{i}\int_0^{i\infty} \exp(\mu(e^{it}-1))dt$$

(as noted above, we may choose the path of integration anywhere in the upper half of the complex plane)

$$\begin{aligned} &= \frac{e^{-\mu}}{i}\int_0^{\infty} \exp(\mu e^{-s})dt \\ &= e^{-\mu}\lim_{L\to\infty}\int_0^{L} \exp(\mu e^{-s})dt \\ &= e^{-\mu}\lim_{L\to\infty}\int_{-\mu}^{-\mu\exp(L)} e^{-t}\frac{dt}{t} \\ &= e^{-\mu}\lim_{L\to\infty}\left(Ei(\mu) - Ei(\mu e^{-L})\right) \end{aligned}$$

$$
\begin{aligned}
&= e^{-\mu} \lim_{L\to\infty} \left(Ei(\mu) - \gamma - \ln(\mu e^{-L}) + \mu e^{-L} \right) \\
&= e^{-\mu} \lim_{L\to\infty} \left(Ei(\mu) - \gamma - \mu + L \right)
\end{aligned}
$$

providing that we choose L such that $\lim_{L\to\infty} \exp(-L) = 0$. This expression diverges, as you might expect from the Cauchy nature of its Gaussian limit. However, if we consider not a Poisson distribution but a modified Poisson where the probability of 0 events vanishes (as makes sense if we are considering photon statistics for a detected object; 0 events corresponds to a non-detection), the characteristic function is modified to $(\exp(\mu e^{it}) - 1)/(\exp(\mu) - 1)$, and we find that

$$
\begin{aligned}
\langle 1/x \rangle &= \frac{1}{e^{\mu} - 1} \left(Ei(\mu) - \gamma - \mu \right) \\
&\approx \frac{1}{\mu} \left(1 + \frac{1}{\mu} + \frac{2}{\mu^2} \right);
\end{aligned}
$$

the discrepancy with the Gaussian result

$$
\frac{1}{\mu} \left(1 + \frac{1}{\mu} + \frac{3}{\mu^2} \right);
$$

is presumably due to the approximation $P(\mu) \approx N(\mu, \mu)$.