

UC Irvine

UC Irvine Previously Published Works

Title

Astronomical incentives to improve angular resolution

Permalink

<https://escholarship.org/uc/item/2sz3q5zp>

Author

Trimble, V

Publication Date

1999

Copyright Information

This work is made available under the terms of a Creative Commons Attribution License, available at <https://creativecommons.org/licenses/by/4.0/>

Peer reviewed

Astronomical Incentives To Improve Angular Resolution

Virginia Trimble

Physics Dept., Univ. of California, Irvine, CA 92697

&

Astronomy Dept., University of Maryland, College Park, MD 20742

Abstract.

Higher angular resolution buys you not only sharper images but also improved signal-to-noise, spectroscopy, and so forth. Thus, in principle, nearly everything that optical and infrared astronomers do can be done better. An overview is provided of examples, with details of a few that happen to appeal to the author.

1. Introduction

The visible and near infrared seeing disks at major observatories range in size from 0.3'' (Mauna Kea at its best, with perhaps a bit of poetic license) up to 2-3'' (many of the older, sea-level sites). Three ways have been devised to cheat Mother Nature out of the limits so imposed: (a) interferometry, including speckle techniques, (b) adaptive optics, and (c) space-based systems, of which the Next Generation Space Telescope is an example.

Each has its pros and cons. Interferometry has been described (by one of its practitioners) as "like being led blindfolded up to the top of a mountain by a single path, and then being required to describe the entire mountain and its Fourier transform". Adaptive optics is a bit like trying to make the wind blow back the way it came before it gets to you. And the launch of NGST will resemble throwing a multi-ton butterfly off the earth and then expecting it to spread its wings and be all things to all PIs. No attempt is made here to distinguish which of the three approaches is most likely to succeed in any particular application.

2. What You (Might) Get

The first obvious reward of shrinking a seeing disk is sharper images with finer detail. But this is not the whole story. Because less sky (and sometimes less detector) will contribute noise photons and counts, you are rewarded with higher signal to noise ratio for a fixed combination of source intensity and exposure time. This means doing astronomy "fainter, further, and faster". There are noise-

limited cases with an almost exact tradeoff between image diameter and telescope diameter (a sobering thought when you ask what comes after a 10 meter).

Third is a coronagraphic effect, enabling detection and study of stuff fainter and closer than you could otherwise handle in the vicinity of bright objects (though not perhaps by quite as much as you would expect, 3^m fainter at $1''$ is a typical limit).

All three of these advantages carry across to spectroscopy and polarimetry. Thus you can expect to get spectra and polarization of finer details, closer environments, and, through enhanced S/N, of fainter, further, and faster astronomical sources. In addition, the smaller images allow you to use narrower slits, without loss of photons, leading again to reduced noise and permitting resolutions of 10^5 and beyond.

Fifth, being able to record a given image or spectrum in less time permits study of more rapidly varying objects. This has not recently been a strong driver (hence the choice of the High Speed Photometer as the HST instrument to sacrifice for COSTAR), but a handful of applications come to mind. Sixth, there are synergies between interferometry and adaptive optics in both directions: optimal application of interferometry with mirrors of more than about 1 meter requires AO, and you can do a better job of AO (especially tip-tilt correction) if you keep track of speckles.

Picking a “reference level of improvement” of $0.03''$ angular resolution, we can expect to be able to say something about objects with distances and sizes of:

- 10 AU – 225 km (*e.g.* the five largest islands on earth, from Uranus)
- 10 pc – 0.3 AU (all sorts of proto-star stuff and binaries)
- 10 kpc – 300 AU = 0.0015 pc (main sequence stars as crowded as 10^4 pc^{-2} in the centers of globular clusters)
- 10 Mpc – 1.5 pc (globular cluster cores)
- 1000 Mpc – 150 pc (width of spiral arms)

3. The Big Questions

We all have our favorites, and most of the lists will look a good deal like the goals of the NASA “Origins” program (illustrating the well-known principle that a large, prestigious committee can’t do these things almost as well as you or I could on our own). Mine includes:

- How far has organic chemistry gone off the earth? This includes complex interstellar molecules (to which improved resolution may be important) and amino acids (*etc.* in meteorites (to which it is probably not). I was particularly struck by the recent confirmation of the excess of left-handed amino acids in Murchison, which occurs in molecules that terrestrial living creatures don’t use and in association with anomalous nitrogen isotope ratios, indicating that contamination is not the cause.

- What causes the initial mass function (number of stars vs. mass at formation)? At least four contributors have been suggested: (1) sizes of cores in molecular clouds (big stars, for instance, seem to form only in big, dense clouds), (2) onset of photodissociation (as in the HST “pillars of creation” image and proplyds), (3) opening up of the bipolar outflow angle at the onset of deuterium burning, convection, and dynamo field generation (this is one of the classic four stages of star formation in the scenario of Shu and his colleagues), and (4) processes involving binary-and multiple cores and stars. This is, of course, just part of the general problem of star formation, but one that sounds particularly vulnerable to studies with improved angular resolution.
- How did galaxies form? This is, I think, the single most important unsolved problem in modern astrophysics. It has, naturally, several pieces: (1) first and foremost, how do you make the matter lumpy on scales from planets to superclusters while leaving the background radiation smooth to a few parts in 10^5 ? (2) what is the nature and role of dark matter? (3) is the dominant process top down (from large scales to small) or bottom up (hierarchical clustering)? (4.) to what extent are elliptical galaxies the products of mergers of disk galaxies? (5) which came first, the chicken bulge or black hole egg? (6) what was the original spectrum of the primordial fluctuations that grow into large scale structure? and (7) what part, if any, do topological defects play in seeding the process? Most of these are interrelated to the point of incest, and sharper images *etc.* should help with many.
- What is the relationship between the dynamical and chemical evolution of galaxies? I call this one “the curse of the variable parameter”, because current models treat the initial mass function, star formation rate (*vs.* redshift, time, metallicity, ...), inflow and outflow of gas with or without metals, mergers, breakups, and so forth as arbitrarily adjustable quantities. Gamow’s claim that “with five parameters you can fit an elephant” proves to be true, without in any way guaranteeing that your-fit actually represents the way nature did it.

Now, all you have to do is match the capabilities of the previous section to the questions and write your proposal. If, however, you are planning to build focal plane instrumentation as part of the deal, I urge you to read Judith Cohen’s letter in the June 1998 AAS newsletter first. The activity is not always as well rewarded as it should be.

4. The Confused Astronomy

This probably describes many of us much of the time but confusion is also a technical term, meaning that you are trying to resolve more than about one source per 25 beamwidths (or point spread functions) and are, therefore, not really counting things at all, but-rather measuring fluctuations in surface brightness or (to a radio astronomer) $P(D)$, the probability of a given-size deflection. The inevitable result in any context where faint things are commoner than bright

ones is an artificial steepening of your source counts, $N(m_v)$ or $\log N - \log S$. A classic example is the Second Cambridge-Catalogue of radio sources, for which $N(S)$ seemed to be a power law of slope -3.2 . This is far enough from the homogeneous, Euclidean prediction of -1.5 that its discoverers claimed that it absolutely falsified Steady State cosmology. But most of the faint sources were not real ones; the correct slope is now known to be about -1.8 for bright sources at meter wavelengths and Steady State is now falsified by many additional observations. Two effects contribute to steepening slopes in confused fields. The obvious one is blending of two or more things below your detection limit to look like one above the limit. The other is a little more difficult to explain and is a sort of photometric Lutz-Kelker correction. The ordinary Lutz-Kelker correction applies to measuring parallaxes and says that you will always find a group of objects to be closer to you than they really are because there are more small parallaxes that are made to look big by observational error than there are large parallaxes that can be made to look small. In the photometric case, suppose you are measuring a bunch of things that register at nine counts (or 9 mJy) each, with an error or uncertainty of ± 3 counts. Some of them are really 12-photon sources with fewer noise counts than average, and some are really 6-photon sources with more noise than average. But the enhanced 6s will outnumber the depressed 12s by a factor of two or so (in the usual homogeneous, isotropic case,) and you will think there are more 9-photon sources than there really are, ending up with too steep a slope for $N(m_v)$. This photometric Lutz-Kelker correction rears its statistical head for S/N less than about four, and is not to be confused with either Malmquist bias or the Scott effect, which have to do with the representation of intrinsically bright things in a magnitude-limited sample.

The effects of confusion and the benefits of reducing it show up in star-clusters and the bulge and center of our own Milky Way, in all other galaxies at some level of star brightness (whether main sequence or luminous blue variables), for dwarf spheroidals in clusters of galaxies, and any time you try to look too deep. For instance, galaxies are only about $1''$ apart at $K=27$, and the Hubble Deep Field would be confused if you tried to duplicate the exposure from the ground. A factor of two improvement in angular resolution means four times as many beam areas in the field. Thus you can go 1.5^m fainter before confusion again sets in. The following sections explore specific examples, in the usual Astronomy 101 order, from comets to cosmology.

5. Solar System

My private favorite example is that dinosaurs with an operating interferometer or AO system could have resolved surface features on the approaching asteroid about a week before it wiped them out. A more serious example is monitoring of hot spots on Io in the infrared, new ones of which signal enhanced volcanic activity. Io is the only body known to be volcanically more active than earth. It has a molten zone because of tidal stretching and compression in its eccentric orbit around Jupiter. With long-term information about the appearance of new hot spots one could, for instance, try to find out whether activity is triggered

at maximum compression or stretching, in the way that new and full moons perhaps act for terrestrial earthquakes.

Other interesting solar system topics include (a) measuring the size, shape, and albedo of cometary nuclei and how they depend on the age and orbit of the comets (we currently have data on about two, and they are not the same) (b) catching new jets as they first appear in cometary comae and associating them with particular spots on the nucleus, and (c) resolving near-Earth asteroids to get sizes, shapes, albedos (composition), and rotation periods and to settle the issue of the reality of companions, all as tracers of the origin of asteroids and their collision histories. One would also like to follow the “weather” on the outer planets through changes in cloud structure, but their surfaces are awkwardly large for most system designs.

6. Astrometry

Nothing was said about this during the oral presentation, but a participant asked about it afterwards. At the moment, I am sufficiently confused by the inconsistencies in results supposedly derived from HIPPARCOS (including my own “1982 proposal” topics) that I am not sure where astrometry is supposed to be going. The fields of view of interferometric and AO systems are not well matched to some kinds of astrometric problems in any case, though clearly small parallaxes and proper motions in crowded fields would be a good bet. One area where higher angular resolution is clearly needed is in registering radio and optical coordinate systems, where it is essential to be sure that the (relatively few) sources in common are points at both wavelengths to within the desired accuracy, or, if they are not, that at least you are registering core-with-core, jet-with-jet, or blobby disk with blobby disk.

7. Stars

Let’s start with one that is not the answer to any fundamental question. A few percent of single white dwarfs have magnetic fields in excess of 10^8 G. Their rotation periods are measured from variability of their linear and circular polarization and of the shapes of their Paschen-Back lines. Known ones range from 1.65 hours to about two weeks. For the best known example, Minkowski’s star, (Grw +70° 8247), and about nine others, the period is either greater than a century or less than about 15 minutes. Periods of a few seconds to a few minutes are perfectly possible, but finding them or ruling them out requires being able to do spectropolarimetry on rather faint stars very quickly. Another case where the reduction in necessary exposure time that comes with improved S/N would be useful is in studying the reprocessing of X-rays into visible light in cataclysmic variables and X-ray binaries.

The problem of deblending gravitational lensing events due to MACHOs (object in the galactic halo) is potentially “fundamental”, since it bears on the nature of dark matter. What you see is a cusp-like brightening of a background star when anything of $10^{-3} - 10^{+3}$ solar masses passes in front. The amplitude of the brightening and the duration of the event tell you the mass of the lens (though only statistically). But if the background star is really a blend, only

part of it will be amplified, and you will conclude both that the amplification factor is smaller and that the duration is shorter than is really the case. This leads to too small an estimated mass for the lens, and so to the wrong conclusion about what sort of object in the galactic halo is contributing to the dark matter.

Perhaps the most obvious stellar case is identification and determination of orbit parameters for binary stars that are too close to be resolved and studied at 1" but too small in velocity amplitude to show up easily as spectroscopic systems. The peak in numbers of systems *vs.* period or semi-major axis seems to fall in this difficult range, so that improvements will matter for our understanding of binary formation. Some other stellar problems where sharper images, higher resolution spectroscopy, *etc.* can help are (1) measuring weak, blended lines of important elements in distant stars, sometimes in clusters, where lithium in all kinds of stars and thorium in old ones are particularly important in the great scheme of things, (2) demonstrating that what you think is the most massive, brightest star at any given instant is always really a cluster, successive improved resolutions of R 136a being a well-known example, and (3) identifying surface features (spots, chemical anomalies, magnetic areas, *etc.*) on cataclysmic variable stars, Am and Ap stars, and their ilk through Doppler tomography, time-resolved polarimetry, and so forth. Another issue that may or may not be fundamental is the amount of turbulence and large-scale magnetic field at the surface of normal stars.

Both can be extracted from spectra, but only if you have $R \geq 10^5$ for many lines of different strengths, Landé g factors, atomic masses, and so forth. Closely related is the problem of finding correct values for rotation periods and magnetic cycle periods (derived from line core emission in Ca II H and K and such) for stars in clusters, where age and composition can be removed from the set of variables – another crowded field problem. In both contexts one is talking about velocities of 1 km/sec and fields of a kilogauss or so.

8. Interstellar Medium and Star Formation

This is territory where the near infrared, where AO and interferometry are a little easier, is the (or at least, a) natural place to look. Things one might do in the ISM itself are to resolve weak, narrow lines, probe fine structure by looking at absorption features in spectra of stars very close together on the sky, and (perhaps most interesting) find out how depletion works in dark clouds by finding the absorption lines in the greatly-reddened and dimmed stars you see where $A_v = 4$. For the newly forming stars themselves, you can hope to resolve disks and jets and learn something about their density and velocity structures as a function of the mass of the underlying star, its own properties, and presence of companions. For instance, are weak lined T Tauri stars really systematically older than the classic ones? Sharper images and more dynamic range mean being able to extend the census of binary systems in star formation regions. This is of some interest because the numbers already found at moderate resolution are comparable with the total binary incidence among main sequence stars. The incidence of triples and whether it is larger or smaller than for main sequence stars can also be studied better.

9. Star Clusters

These are the quintessential crowded fields, especially the globular clusters, even in our own galaxy. Thus one can expect to improve our knowledge of the initial mass function at low masses and how it depends on cluster age, metallicity, location in the galaxy, and whatever else matters, and of the incidence of eclipsing, spectroscopic and astrometric binaries. Mass segregation (the dependence of $N(M)$ on location in a cluster) has already been seen many places where ages are smaller than expected dynamically timescales. Thus the question of whether it is an “inherited” or an “acquired” characteristic arises, and good data on $N(M, r)$ in many clusters of different ages and densities can provide the answer.

Other cluster issues where there is potential for progress include:

(1) resolving and counting individual asymptotic giant branch stars in globular clusters in M 31, (2) deciding whether the clusters now forming in mergers and interactions like that of NGC 1275, will end up as globulars (you need a good velocity dispersion to get the masses), (3) distinguishing between cuspy and thermal cluster cores as far as Virgo and so deciding how the probability of core collapse depends on the mass of the host galaxy, the location of the cluster in it, and so forth, (4) identifying variables on the main sequence and getting radial velocities and proper motions to be sure they are members, and (5) if you can push your technique into the 3.5μ window, looking for brown dwarfs in older clusters, maybe even globulars.

Finally, another task that bears on the big questions is that of finding the oldest white dwarfs in globular clusters. Their colors and luminosities, fitted to models of cooling, will provide an independent lower limit to the age of the universe. This is not quite an HST problem because you need very clean colors for faint stars.

10. The Milky Way

Our central black hole is currently probed via the locations and motions of giant stars in the cluster around it. It would be nice to extend down to main sequence stars, both to learn more about the central gravitational potential and to find out what kinds of stars form and live in such an extreme environment. The metallicity of stars in the central galactic bulge remains in some dispute. This too is a “faint, crowded, infrared” problem, and the answer is important not just for understanding galactic chemical evolution but also for estimating how many stars might reasonably host terrestrial planets, made of heavy elements.

11. Galaxies

One can, of course, think both of things to be done about stuff inside other galaxies and of things that concern populations of galaxies. Working, again, crudely, from near to far, I think it would be interesting to (a) get a good census of sizes and luminosities of HII regions for 9 Virgo galaxies and beyond and find out whether $N(L)$ really is universal as currently claimed, (b) get velocity dispersions for lots of star formation regions and so dynamical masses

as a handle on the H:CO conversion factor and how it depends on everything, (c) look for and measure masses of central black holes in lots more galaxies of different types (this means following changes in velocity field in very close to the center) and see whether the correlation of black hole mass with mass of the bulge component holds up, and (d) find the globular cluster populations of more distant galaxies and perhaps begin to say something observational about how their numbers and properties depend on their parent galaxies when the galaxies aren't quite like $z = 0$ ones. Among the many things one might want to do with galaxies as wholes are (a) deep counts at 2.2 and 3.5μ to find ones where star formation turned off long ago and to see how much evolution of various kinds has occurred in the older populations (both stellar and galactic), (b) find faint blue galaxies and dwarf spheroidals in both the field and clusters at moderate redshift (the "big question" here is whether these types and others are part of an evolutionary sequence, perhaps related to some sort of QSO absorption line cloud type), and (c) to correlate morphology with spectra, colors, luminosities, rotation curves and so forth, at redshifts of 1 and more (since even at $z = 0.3$ one finds rather considerable differences from $z = 0$, approximately summarized as the Butcher-Oemler effect).

12. Quasars, Active Galaxies, and Star Bursts

Another topic where "fainter, closer, infraredder" is obviously relevant is that of QSO host galaxies at redshifts of three or more.

Very nearby ones seem to live mostly in small groups and to conform at least crudely to the rule radio loud = elliptical, radio quiet disk. By $z = 0.5 - 1$, they are turning up in rich clusters (and the rule is probably breaking down). What happens beyond that, especially at redshifts where clustering is not yet much developed? It has been claimed that merger $\rightarrow$ star burst $\rightarrow$ QSO is an evolutionary sequence. Is this true? One handle would be to look for the correlation of the amount of nuclear (AGN) light that contributes in polarized or reflected light with morphology, again preferably at moderate redshift and beyond.

What else might you resolve? Probably the narrow line region, at least in nearby Seyferts, thereby finding out whether it is an extended accretion disk, a sort of generic haze made of wind losses from lots of stars, or something else. Certainly one can push on core-jet structure (though not immediately to the $0.001''$ resolution available from radio VLBI!). Why bother? Well, lots of theorists are interested in magnetic collimation and such. But also showing that optical and radio polarization line up very precisely can be used to put limits on how photons propagate over large distances and on whether the universe has any innate chirality. I feel attached to this because of a 1997 paper that claimed to have found chirality, quickly contradicted by people with access to better data, but, also in conflict with an upper limit — that Maurice Goldhaber and I had published a little earlier, unfortunately (perhaps the word is foolishly) in a non-prestigious journal.

QSOs also function as probes of intervening material. Some of the things you can do with fainter sources and higher spectral resolution include resolving the Lyman-alpha forest to look for the much weaker Deuterium line and to decide

whether there is any continuous absorption (Gunn-Peterson effect) between the discrete lines and also looking for absorption lines close to the emission redshift in intrinsically rather faint QSOs to decide how much intergalactic ultraviolet flux is contributing to ionization of the clouds. It has been claimed that observations of QSO absorption-lines were the single largest absorber of large telescope time in the 1980s. One can imagine this again being true for time on large interferometer arrays (but not immediately, because they won't go quite faint enough).

13. Cosmology, Distance Scales, etc.

An obvious set of applications is the calibration, extension, and verification of assorted distance indicators. For instance, (1) for the luminosity function of planetary nebulae, one can both get $N(L)$ for many more, and more distant, galaxies and use improved-spectra to try to figure out the nature of the emission line objects that are too bright to be PNe and distort the upper end of $N(L)$, (2) for the globular cluster luminosity function, one can try to reach the turn-over at $M_v = -6.5$ (or so) in enough galaxies to decide what the details of $N(L)$ are correlated with among possible properties of the host galaxies (sadly, one does not reach the peak at large enough redshift to expect to see evolutionary changes), (3) for the Tully-Fisher relation for disk galaxies and the fundamental plane for ellipticals, one can reach $z = 0.1$ to 0.5 , directly coupling to Type Ia supernovae in the regime where they are being used to get cosmological parameters beyond H_0 , and (4) for the indicator that uses fluctuations in surface brightness as the standard candle, you can both reach out to the Virgo cluster and beyond (this has already been done for a few galaxies with HST) and, perhaps more important, identify some of the individual stars that are contributing to SB fluctuations in closer galaxies; this is of importance because, as Tonry said a couple of years ago, "Either H_0 is greater than 80 km/sec/Mpc, or I am working on a very interesting problem in stellar populations", and it begins to seem as if the second alternative must be the right one. The dependence of average surface brightness on redshift also constitutes a cosmological test of a slightly different kind. It scales as $(1+z)^{-1}$ for tired light, $(1+z)^{-2}$ for expansion of galaxies and clusters into existing space, and $(1+z)^{-4}$ for general relativistic expansion of the space-time itself (which is what we all believe, even if we don't quite say it that way as a rule). The "tired light" case has already been ruled out. Distinguishing the latter two means working at redshift 0.3 or more and knowing enough about the properties of the galaxy you are looking at to be able to compare it with the right $z = 0$ ones.

Supernovae are important to the great scheme of things both as distance indicators and as the main contributors to galactic chemical evolution. Spectra for events at $z \geq 1$ can be used (a) to test the extent to which the properties of Type Ia supernovae depend on their hosts, progenitors, *etc.*, (b) to get rates of the different types *vs.* time and type of host galaxy (as a check on the usual claim that the metals in low-metallicity stars come mostly from core collapse, Type II events), and (c) to see whether the total production of heavy elements adds up to what we think we need at various redshifts to match colors of galaxies, metals in the gas in X-ray clusters, and the compositions of QSO absorption clouds of various types.

Of course you can also use your improved sensitivity and angular resolution to plot light curves for Cepheids in the Coma cluster or thereabouts. But I hope very much no observatory will declare this to be a “key project”, or we may end up seeing Paper MCMXLVIII of the series.

14. With Three You Get Eggroll

You probably feel that keeping an interferometric array or an adaptive optics system working is already as much as any team of astronomers can be expected to do. True enough, but, since I am not the one who has to do it, I did consider briefly what good it would do to stick yet another widget into the light path. For instance, adding a Fabry-Perot spectrometer would enable you to map the turbulence in the Crab Nebula in enough detail to see how it correlates with the structure of the magnetic field. Or you could stick in an objective prism or grism and find the individual Wolf-Rayet stars in Wolf-Rayet galaxies (which are currently identified only by their integrated spectra). And, undoubtedly, if you think about it, other applications of greater importance to the “big questions”.

Acknowledgments. I am grateful to Gary Chanan for sharing his copy of the unpublished volume on Adaptive Optics for the Keck Telescopes (of which he was also one of the authors and editors) as well as his knowledge of how to do difficult kinds of astronomy, and to the organizers for the invitation to give this presentation, which forced me to examine previously-ignored corners of astronomical parameter space.