Redundant spacing calibration for actual atmospheric turbulence

Wilson McKellar, Imperial College, London

Abstract

A description of current experimental work will be given. This repeats a Redundant Spacings Calibration (RSC) experiment carried out by Greeenaway (et. al.) [+], but incorporating real atmospheric turbulence. A nine aperture redundant array produces a system of fringes (modulated by an aperture diffraction envelope), using light from an expanded HeNe laser beam, across approximately 180m of atmosphere. The RSC technique is implemented and a point spread function reconstructed from the corrected fringe phases. Success is gauged on the Strehl ratio of the corrected image.

REDUNDANT SPACINGS CALIBRATION IN THE PRESENCE OF REAL ATMOSPHERE

Wilson McKellar

Imperial College of Science
Technology and Medicine

Applied OPTICS

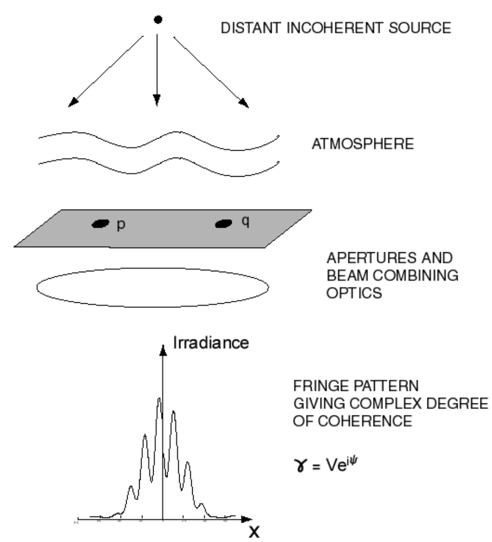
Statement of the Problem

- Require true object phase, ϕ_{pq}, from measured phase, ψ_{pq}, with phase errors e_1 to e_N over each aperture. Measured phase is now given by

$$\psi_{pq} = \phi_{pq} + e_p - e_q$$

or in matrix form

$$\Psi = A\Phi$$



Applied OPTICS

Redundant Spacings Calibration

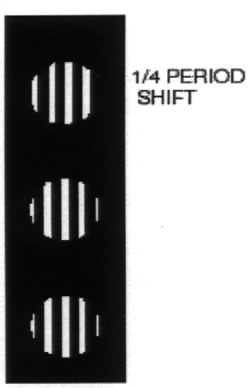
$$\Psi = A\Phi$$

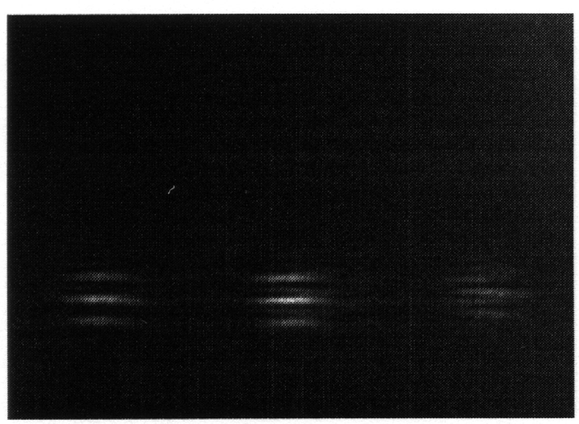
- N aperture array gives $M = N(N-1)/2$ corrupted phase measurements. Matrix A is rank deficient by N.
- For 2 dimensional array one e_p and two values of ϕ may be arbitrarily set to zero.
- N-3 independent redundancy conditions bring A to full rank. The pseudo inverse of A produces a unique least squares. (Only true for critical redundancy).

Applied OPTICS

Snapshot technique used by Greenaway et. al.

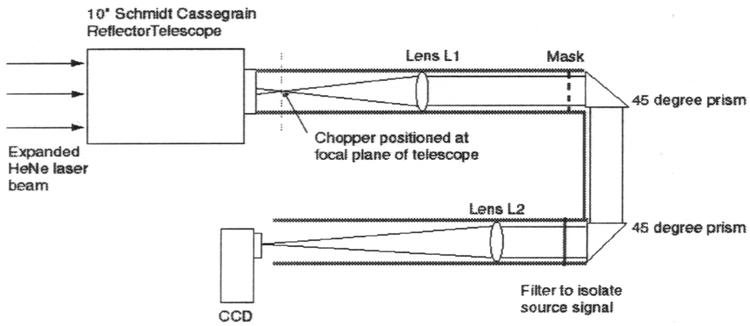
- Redundant baselines distinguished at the same position in the spatial frequency plane using a dislocated grating.

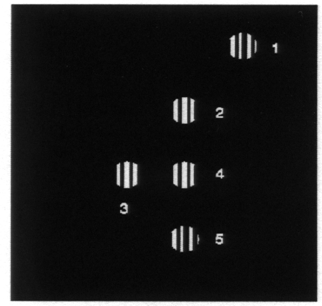




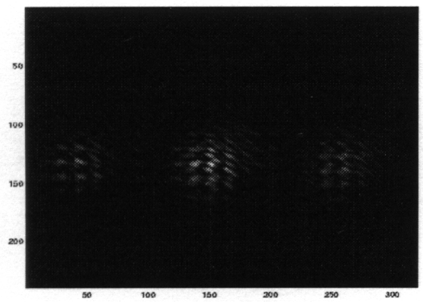
Applied OPTICS

Experiment at Imperial





REDUNDANT MASK



CCD Image showing the array fringes contained within the grating diffraction orders.

Applied OPTICS

ON THE USE OF LUNAR OCCULTATION AND INTERFEROMETRIC OBSERVATIONS IN COMBINATION FOR INVESTIGATIONS OF STARS

Eugene M. Trunkovsky

Sternberg Astronomical Institute, Moscow State University,
Universitetskij prospect 13, Moscow, 119899, Russia
e-mail: tem@sai.msu.ru

Abstract

Interferometric and lunar occultation observations of stars are both the effective techniques which allow us to reach high angular resolution at a level of 1 milliarcsecond and better both in the visual and in the near-IR spectral ranges, and therefore permit to investigate directly the spatial structure of various stellar sources and to determine on this basis, apart from indirect (theoretical) considerations, a principal physical characteristics of the objects under study. By now a rather great number of direct measurements of the angular diameters of stars having different spectral types and luminosities has been carried out by these two techniques; typical values of the measured diameters are as small as a few milliarcseconds; the minimal reliably determined ones are smaller than the value of $0''.001$. Also, numerous close binary and multiple stars with the angular distances between the components ranging from $0''.5$ down to $0''.002$ have been discovered and studied. On this basis the most important physical parameters of the objects investigated were determined such as the linear radii and effective temperatures in case of a single stars, and the orbital elements, luminosity and mass ratios of the components in case of a close binary or multiple stellar systems.

At the same time, there are rather many cases, when the results obtained from different observations or by different ways prove to be inconsistent. The discordances may be caused by previously unknown close binarity or multiplicity of the stars under study, or by presence of some circumstellar material (as extended envelope or atmosphere, disk-like structure etc.) around the star itself, or by variations of the object's dimensions with time (e.g. pulsations of a star). In such cases it is very desirable to get independently a high-quality data on the actual spatial structure of the objects investigated and to check a previously obtained results. Since an opportunities of realizing a repeated research of the angular structure of the sources of interest by means of lunar occultation observations are strictly limited due to a strong dependence of observing chance on the path of lunar motion, an importance of interferometric investigation of such objects with use of a new generation of a ground based and space interferometric systems is very great. The two discussed techniques are thus complementary and should be used in combination.

MAXIMUM LIKELIHOOD AND BAYESIAN METHODS FOR SYNTHESIS IMAGING

JORGE NÚÑEZ
UNIVERSITY OF BARCELONA

PROBLEM: ESTIMATE SKY BRIGHTNESS FROM VISIBILITY

$$\tilde{V}(u,v) = \iint_{-\infty}^{\infty} a(x,y)\, e^{-2\pi i (ux+vy)}\, dx\,dy = FT[a(x,y)]$$

$$a(x,y) = \iint_{-\infty}^{\infty} \tilde{V}(u,v)\, e^{+2\pi i (ux+vy)}\, du\,dv = IFT[\tilde{V}(u,v)]$$

$\tilde{V}(u,v)$ = COMPLEX VISIBILITY

$a(x,y)$ = TRUE TWO-DIMENSIONAL INTENSITY DISTRIBUTION OR "OBJECT" (TRUE MAP)

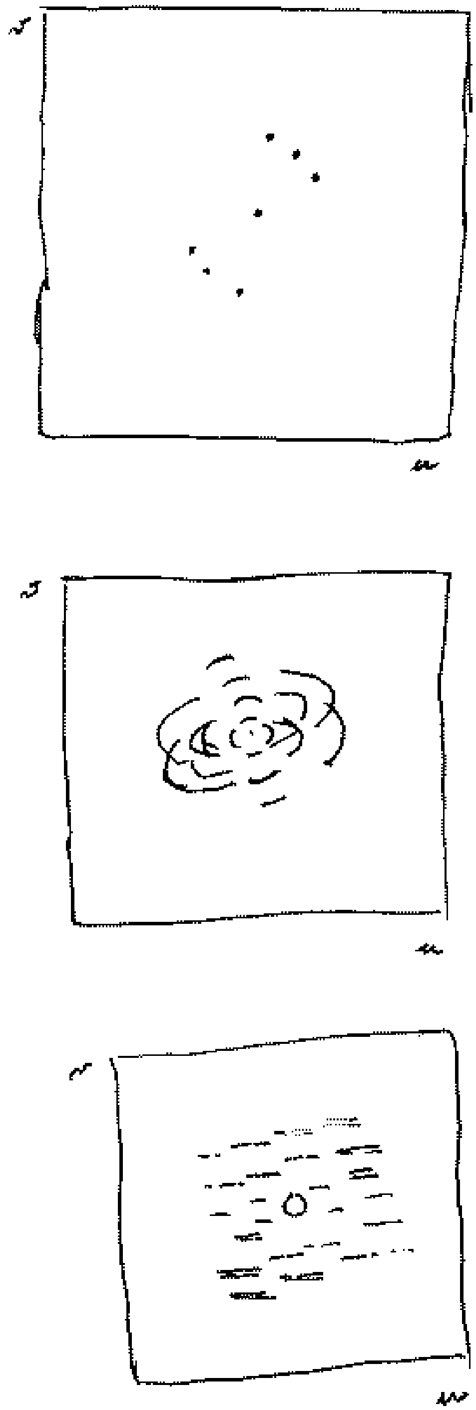
- VISIBILITY $\tilde{V}(u,v)$ IS SAMPLED AT DISCRETE POINTS AND UP TO A MAXIMUM BASELINE. THIS CAN BE DESCRIBED BY A SAMPLING FUNCTION

$$F(u,v) = \sum_{(K)} \delta(u-u_K, v-v_K)$$

(K): "KNOWN" (MEASURED) VISIBILITY

VISIBILITY OFTEN NEEDS DIFFERENT WEIGHTS AT DIFFERENT SAMPLING POINTS. USING A WEIGHTING (TAPER) FUNCTION:

$$F(u,v) = \sum_{(K)} t(u,v)\, \delta(u-u_K, v-v_K)$$



– THE OBSERVED VISIBILITY:

$$\boxed{\tilde{V}_{obs}(u,v) = \tilde{V}(u,v) \cdot F(u,v)}$$

THUS, THE "OBSERVED" INTENSITY OR IMAGE (DIRTY MAP)

$$P(x,y) = \iint_{-\infty}^{\infty} \tilde{V}_{obs}(u,v)\, e^{+2\pi i(ux+vy)}\, du\,dv =$$

$$= \iint_{-\infty}^{\infty} \tilde{V}(u,v)\, F(u,v)\, e^{+2\pi i(ux+vy)}\, du\,dv$$

$$= a(x,y) * F(x,y) \qquad \text{CONVOLUTION}$$

WHERE

$$F(X,Y) = \iint_{-\infty}^{\infty} F(u,v)\, e^{+2\pi i(ux+vy)}\, du\,dv = IFT\left[F(u,v)\right]$$

IS THE POINT SPREAD FUNCTION (DIRTY BEAM)

– THE IMAGING EQUATION

$$\boxed{P(x,y) = a(x,y) * F(x,y)}$$

OPTICAL/IR: IMAGE = OBJECT * PSF

RADIO: DIRTY MAP = TRUE MAP * DIRTY BEAM

– IN FOURIER DOMAIN

$$\boxed{\tilde{V}_{obs}(u,v) = \tilde{a}(u,v) \cdot F(u,v)}$$

$\tilde{a}(u,v) = FT\left[\text{ESTIMATED OBJECT}\right] \simeq \tilde{V}(u,v)$

FURTHERMORE

- $a(x,y)$, $P(x,y)$ ARE STATISTICAL REALIZATIONS (POISSON, GAUSS..) OF THE "TRUE" INTENSITY DISTRIBUTION

- DATA ARE NOISY (IN PARTICULAR THE PHASE OF THE OBSERVED COMPLEX VISIBILITY $\tilde{V}_{obs}(u,v)$)

THE IMAGING EQUATION BECOMES

$$\boxed{\tilde{V}_{obs}(u,v) = \tilde{a}(u,v)\, F(u,v) + \tilde{N}(u,v)}$$

- NOISE CORRUPTS THE HIGH-ORDER-FREQUENCY

- THE IMAGE RECONSTRUCTION IS AN ILL-POSED INVERSION PROBLEM

- THE GOAL OF IMAGE RECONSTRUCTION IN APERTURE SYNTHESIS IS TO EXTRACT FROM THE MEASURED VISIBILITY ($\tilde{V}_{obs}$) AN APPROXIMATION TO THE TRUE OBJECT ($a(x,y)$) WITH REDUCED RIPPLE AND POSSIBLY IMPROVED RESOLUTION, MAKING ALLOWANCE FOR NOISE. THIS INVOLVE SOME FORM OF INTERPOLATION AND EXTRAPOLATION IN THE FOURIER DOMAIN.

- A GOOD ALGORITHM SHOUL ALSO GIVE LESS WEIGHT TO HIGH-SPATIAL-FREQUENCY DATA THAT ARE BADLY CORRUPTED BY NOISE AND INSTEAD GIVE MORE IMPORTANCE TO EXTRAPOLATED ~~DATA~~ VALUES OBTAINED FROM BETTER DATA AT LOWER FREQUENCIES.

$$g(x) = \int_0^x f(y)\,dy \Rightarrow f(y) = \left(\frac{dg}{dx}\right)_{x=y}$$

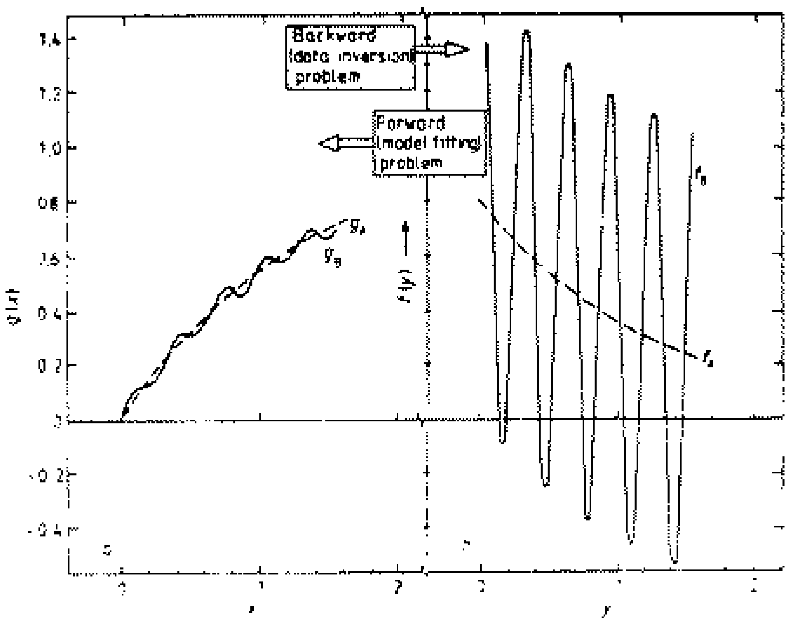


Figure 1.1 Illustration of the properties of the very simple equation (1.9) and its inverse (1.10) for two specific functions $g_A(x) = 1 - \exp(-\alpha x)$ (broken curve) and $g_B(x) = 1 - \exp(-\alpha x) + \beta \sin \omega x$ (full curve) using the numerical values $\alpha = 0.8$, $\beta = 0.04$, $\omega = 20$. Though the functions g_A and g_B (figure 1.1(a)) differ only by a small ($\simeq 5\%$) high frequency ripple in function space they are much more widely separated in derivative space. Consequently the corresponding solutions in equation (1.10) of equation (1.9), namely $f_A(y) = \alpha \exp(-\alpha y)$ and $f_B(y) = \alpha \exp(-\alpha y) + \omega\beta \cos \omega y$ differ by more than 100% as shown in figure 1.1(b). Consequently the solution $f(y)$ of the backward (integral inversion) problem is unstable to high frequency changes in $g(x)$. Conversely the forward (model fitting) problem, wherein functions $f(y)$ are compared in terms of their prediction of $g(x)$, is extremely insensitive to large high frequency changes in $f(y)$.

Evidently for a given ripple amplitude β, the problem becomes progressively worse with increasing ω. In particular, if the ripple on $g(x)$ has a saw-tooth form then it has pointwise infinite derivatives and the oscillations of solution $f(y)$ become infinite. (Aternatively stated, a saw-tooth in $g(x)$ contains Fourier components of infinite frequency ω so that the amplitude $\omega\beta$ of the resulting ripple component in $f(y)$ becomes infinite even for arbitrarily small amplitude β.)

RECONSTRUCTION METHODS

- LINEAR METHODS (DIRECT FT, FILTERED BACK-PROJECTION, REGRIDDING AND FFT...) CANNOT GENERATE NON ZERO VALUES AT UNMEASURED FREQUENCIES NOR CAN EXTRAPOLAT INTO NOISE-CONTAMINATED REGIONS OF THE (u,v) PLANE.

- NONLINEAR METHODS (CLEAN OR STATISTICALLY-BASED METHOD: MAXIMUM ENTROPY (MEM), STOPPED MAXIMUM LIKELIHOOD, BAYESIAN,...) GENERATE NON ZERO VALUES FOR UNMEASURED OR VERY NOISY (u,v) POINTS. THE USE "A PRIORI" INFORMATION SUCH AS POSITIVITY OF INTENSITY FINITE SPATIAL EXTENT OF THE SOURCE, SMOOTHNESS, ET A NATURAL EFFECT OF THIS EXTRAPOLATION IS INCREAS RESOLUTION. A MODEST DEGREE OF SUPERRESOLUTION (UP TO A FACTOR OF TWO) WILL OFTEN BE POSSIBLE USING UNINFORMATIVE (POSITIVITY) PRIORS. THUS, IT IS POSSIBLE TO GO BEYOND THE DIFFRACTION LIMIT OF THE INSTRUMENT (FOR EXAMPLE IN A SPACE INTERFEROMETER).

CLEAN : FAVORED BY RADIOASTRONOMERS

ROBUST

CONVERGENT EVEN IN DIFFICULT SITUATIONS

MATHEMATICAL BEHAVIOR NOT WELL UNDERSTOOD

ERROR ANALYSIS DIFFICULT

STATISTICALLY-BASED METHODS

BAYESIAN METHODS

BAYES THEOREM : $P(\vec{a}|\vec{P}) = \dfrac{P(\vec{P}|\vec{a})\,P(\vec{a})}{P(\vec{P})}$

$P(\vec{a}|\vec{P})$ = PROBABILITY THAT OBJECT $\vec{a}$ BE TRUE GIVEN DA

$P(\vec{P}|\vec{a})$ = LIKELIHOOD = PROBABILITY THAT DATA $\vec{P}$ COME FROM OBJECT $\vec{a}$. DESCRIBES THE NOISE IN TH DATA AND ITS POSSIBLE OBJECT DEPENDENCE.

$P(\vec{P})$ = NORMALIZATION CONSTANT

$P(\vec{a})$ = PRIOR PROBABILITY. IT IS KNOWN PRIOR TI OBTAINING THE DATA.

SEVERAL PRIORS ARE POSSIBLE

- GIBBS ENERGY FUNCTIONS

  $P(\vec{a}) \propto e^{-U(\vec{a})}$

- ALOTATION PRIORS

  EX. $P(\vec{a})$ = CONST $0 \leq a < b$

  $P(\vec{a}) = 0$ ELSEWHERE

- POSITIVITY OF INTENSITY

- SOURCES ARE POINTS (CLEAN) OR SMOOTHLY EXTEND OR DISKS ..

- NEGATIVE PRIORS

  EX.: NOT A GALAXY

— ENTROPY

USING ONE-DIMENSIONAL NOTATION (AFTER PIXELIZATION)

$$\boxed{\log \mathcal{P}(a) = -\frac{1}{\Delta a} \sum_i a_i \log \frac{a_i}{Q_i}}$$ (CROSS-ENTROPY

- Q_i IS A DEFFAULT IMAGE (PRIOR ENERGY DISTRIBUTION) OBTAINED FROM OTHER OBSERVATIONS OR ASSUMPTIONS. IF Q_i = CONST. THERE IS NO DEFFAULT (UNINFORMATIVE PRIOR)
- Δa IS AN ADJUSTABLE PARAMETER
- WITHOUT DATA THE SOLUTION IS Q_i
- THE PRACTICAL INTEREST OF USING ENTROPY IS THA THE RESULTING RECONSTRUCTION HAS MINIMAL CONFIGU RATIONAL INFORMATION CONSISTENT WITH THE DATA, SO THERE MUST BE EVIDENCE IN THE DATA FOR ANY STRUCTURE THAT IS SEEN.

LIKELIHOOD

DESCRIBES THE NOISE IN THE DATA

- GAUSSIAN (σ = CONST.)
- GAUSSIAN (σ = VARIABLE)
- POISSON (PULSE COUNTING DEVICES IN LOW LIGHT SITUATION)
- POISSON + GAUSS (CCD)
- ETC.

ALGORITHMS DEPEND STRONGLY ON THE NOISE MODEL

— POISSON

$$\log \underline{P}(P|a) = \log L = \sum_j \left[-(a*F)_j + P_j \log (a*F)_j - \log P_j! \right.$$

GIVES THE POWERFUL RICHARDSON-LUCY ALGORITHM (MAXIMUM LIKELIHOOD) OR "R-L"

$$a_i^{(k+1)} = a_i^{(k)} \sum_j F_{ji} \frac{P_j}{(a^{(k)} * F)_j} = a_i^{(k)} F * \frac{P}{a^{(k)} * F}$$

- THE ITERATIVE ALGORITHM IS POSITIVE BECAUSE THE DATA IS POSITIVE, CAN BE COMPUTED USING FFT
- THE ALGORITHM MUST BE STOPPED BEFORE MAXIMUM. IT HAS SEVERAL IMPROVEMENTS (MULTIPLE CHANNELS)

INCLUDING THE ENTROPY PRIOR

$$a_i^{(k+1)} = K a_i^{(k)} \left[\sum_j F_{ji} \frac{P_j}{(a^{(k)} * F)_j} - \frac{1}{\Delta a} \left(\log \frac{a_i^{(k)}}{Q_i} + 1 \right) + C \right]$$

THE WEIGHT OF THE ENTROPY (Δa) SHOULD BE COMPUTED, IT CAN BE DIFFERENT IN DIFFERENT AREAS

— POISSON + GAUSSIAN

$$\log \underline{P}(P|a) = \log L = \sum_j \left[-\log \sqrt{2\pi}\,\sigma - (a*F)_j + \log \sum_{k=0}^{\infty} \left(e^{-\frac{(k-P_j)^2}{2\sigma^2}} \cdot \frac{(a*F}{k!} \right. \right.$$

SIMILAR ALGORITHM BUT USING IN PLACE OF DATA P_j A BAYESIAN FILTERED VERSION OF THE DATA.

$$P_j' = \frac{\sum_0^{\infty} \left(k\, e^{-\frac{(k-P_j)^2}{2\sigma^2}} \left[(a*F)_j^k / k! \right] \right)}{\sum_0^{\infty} \left(e^{-\frac{(k-P_j)^2}{2\sigma^2}} \left[(a*F)_j^k / k! \right] \right)}$$

— THE ADVENTAGE OF THE STATISTICALLY-BASED METHODS IS THAT STOPPING POINTS OR ENTROPY WEIGHTS CAN BE COMPUTED BY ROBUST METHODS LIKE CROSS-VALIDATION

- MAXIMUM ENTROPY METHOD (MEM) IS SOMEWHAT DIFFE[RENT]
BECAUSE MAXIMIZES THE ENTROPY WITH CONSTRAINTS
(EX: FIXING A VALUE FOR χ^2)

WORK IN PROGRESS FOR SYNTHESIS IMAGING AT BARCE[LONA]

- DATA SHOULD NOT BE MOVED FROM THEIR ORIGIN
- COMPUTE LIKELIHOOD IN FOURIER DOMAIN

A) GAUSSIAN WITH $\sigma =$ CONST

IN PLACE OF $P(P\,|\,a)$ IN OBJECT (x,y) SPACE COMPUTE

$$P(\tilde{V}_{obs}\,|\,a) = \prod_{u,v \in K} \frac{1}{\sqrt{2\pi}\,\sigma}\, e^{-\frac{|\tilde{V}(u,v) - \tilde{a}(u,v)|^2}{2\sigma^2}}$$

USING THE HERMITIAN PROPERTY OF VISIBILITY ($\tilde{V}(u,v) = \tilde{V}^*(-u,-v)$)

TAKING THE MAXIMUM RESPECT $a(x,y)$ USING

$$\frac{\partial \tilde{a}(u,v)}{\partial a(x,y)} = e^{-\frac{2\pi i}{N}(ux+vy)}$$

MAXIMIZE THE LIKELIHOOD $\dfrac{\partial \log P(\tilde{V}_{obs}\,|\,a)}{\partial a(x,y)} = 0$ GIVES

$$\sum_{u,v}^{K} (\tilde{V}(u,v) - \tilde{a}(u,v))\, e^{\frac{2\pi i}{N}(ux+vy)} = \mu\sigma^2$$

USING THE SAMPLING FUNCTION $F(u,v)$ AND IFT

$$\sum_{u,v}^{N-1} (\tilde{V}(u,v) - \tilde{a}(u,v))\, F(u,v)\, e^{\frac{2\pi i}{N}(ux+vy)} = \mu\sigma^2$$

MAXIMUM LIKELIHOOD ALGORITHM

$$\boxed{a^{(k+1)}(x,y) = a^{(k)}(x,y) + \text{IFT}\left(\tilde{V}(u,v)F(u,v) - \tilde{a}(u,v)F(u,v)\right) + \text{CONST}}$$

B) GAUSSIAN WITH σ = VARIABLE BUT INDEPENDENT OF SIGNAL

(EX: $\sigma(\mu,\nu) = |\tilde{V}(\mu,\nu)|^{1/2}$)

SIMILAR ALGORITHM:

$$\boxed{a^{(k+1)}(x,y) = a^{(k)}(x,y) + \mathrm{IFT}\left(\frac{\tilde{V}(\mu,\nu)F(\mu,\nu) - \tilde{a}(\mu,\nu)F(\mu,\nu)}{F\,|\tilde{V}(\mu,\nu)|}\right) + \text{CONST.}}$$

C) GAUSSIAN WITH σ = VARIABLE DEPENDENT OF SIGNAL

CASE $\sigma(\mu,\nu) = |\tilde{a}(\mu,\nu)|^{1/2}$

THE LIKELIHOOD

$$P(\tilde{V}_{obs}\,|\,a) = \prod_{\substack{\mu,\nu\\ \in K}} \frac{1}{\sqrt{2\pi}\,|\tilde{a}(\mu,\nu)|^{1/2}}\, e^{-\frac{|\tilde{V}(\mu,\nu) - a(\mu,\nu)|^2}{2|\tilde{a}(\mu,\nu)|}}$$

DERIVATIVES ARE MORE COMPLICATED BECAUSE OF THE DEPENDENCE

OF $\sigma(\mu,\nu)$ OF $a(x,y)$

$$\frac{\partial \log P(\tilde{V}_{obs}\,|a)}{\partial a(x,y)} = 0 \quad \text{GIVES}$$

$$\sum_{\mu,\nu}^{K} \frac{|\tilde{V}(\mu,\nu)|^2 - |\tilde{a}(\mu,\nu)|^2}{2|\tilde{a}(\mu,\nu)|^3}\, \tilde{a}(\mu,\nu)\, e^{\frac{2\pi i}{N}(\mu x + \nu y)} = \mu$$

USING THE SAMPLING FUNCTION, AND THE IFT,

MAXIMUM LIKELIHOOD ALGORITHM

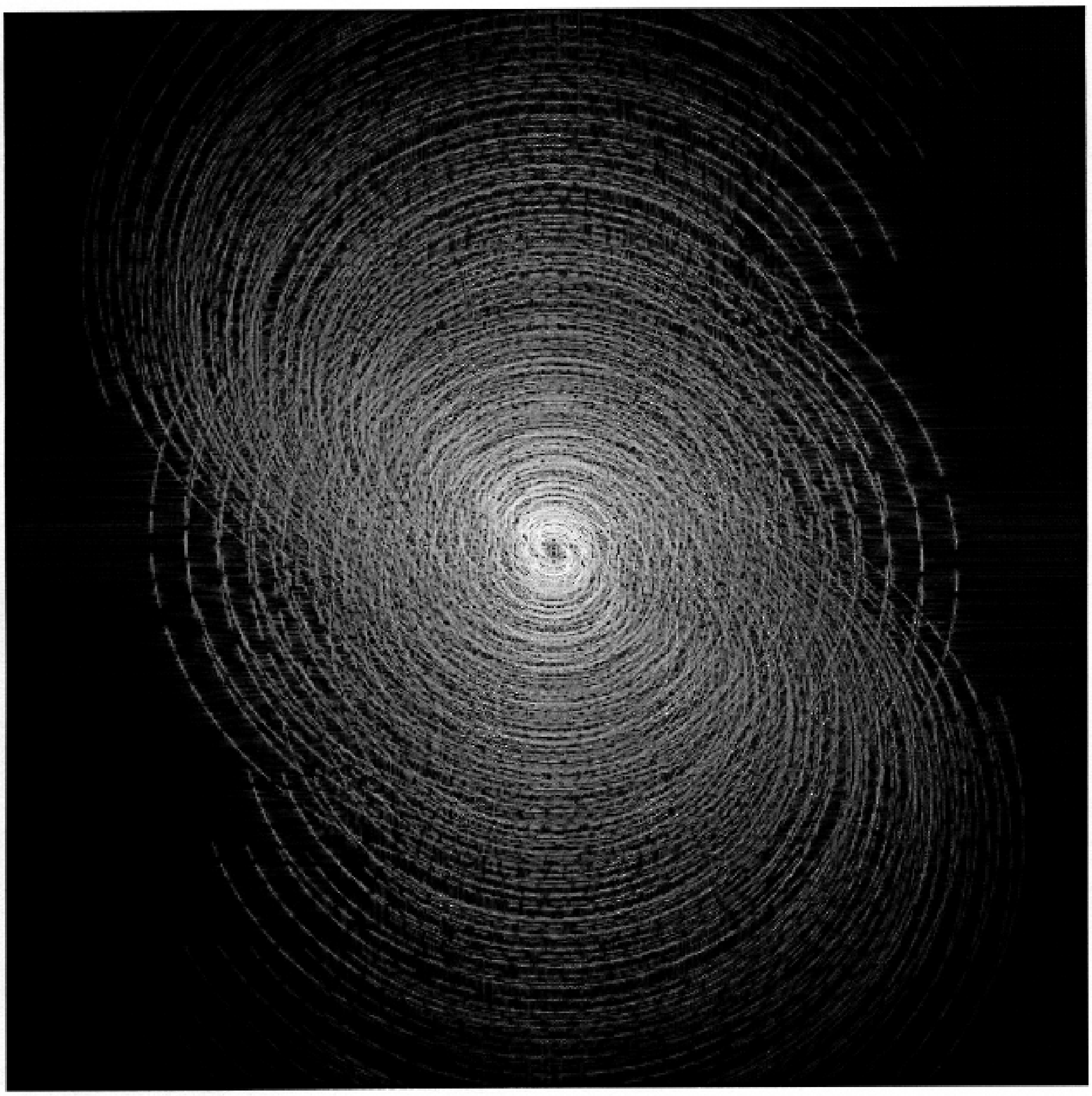
$$\boxed{a^{(k+1)}(x,y) = a^{(k)}(x,y) + \mathrm{IFT}\left[\frac{|\tilde{V}(\mu,\nu)F(\mu,\nu)|^2 - |\tilde{a}(\mu,\nu)F(\mu,\nu)|^2}{2\,|\tilde{a}(\mu,\nu)F(\mu,\nu)|^3}\,\tilde{a}(\mu,\nu)\right]}$$

- ALGORITHMS SHOULD BE STOPPED
- CAN INCORPORATE THE ENTROPY PRIOR

LSI +61° 303

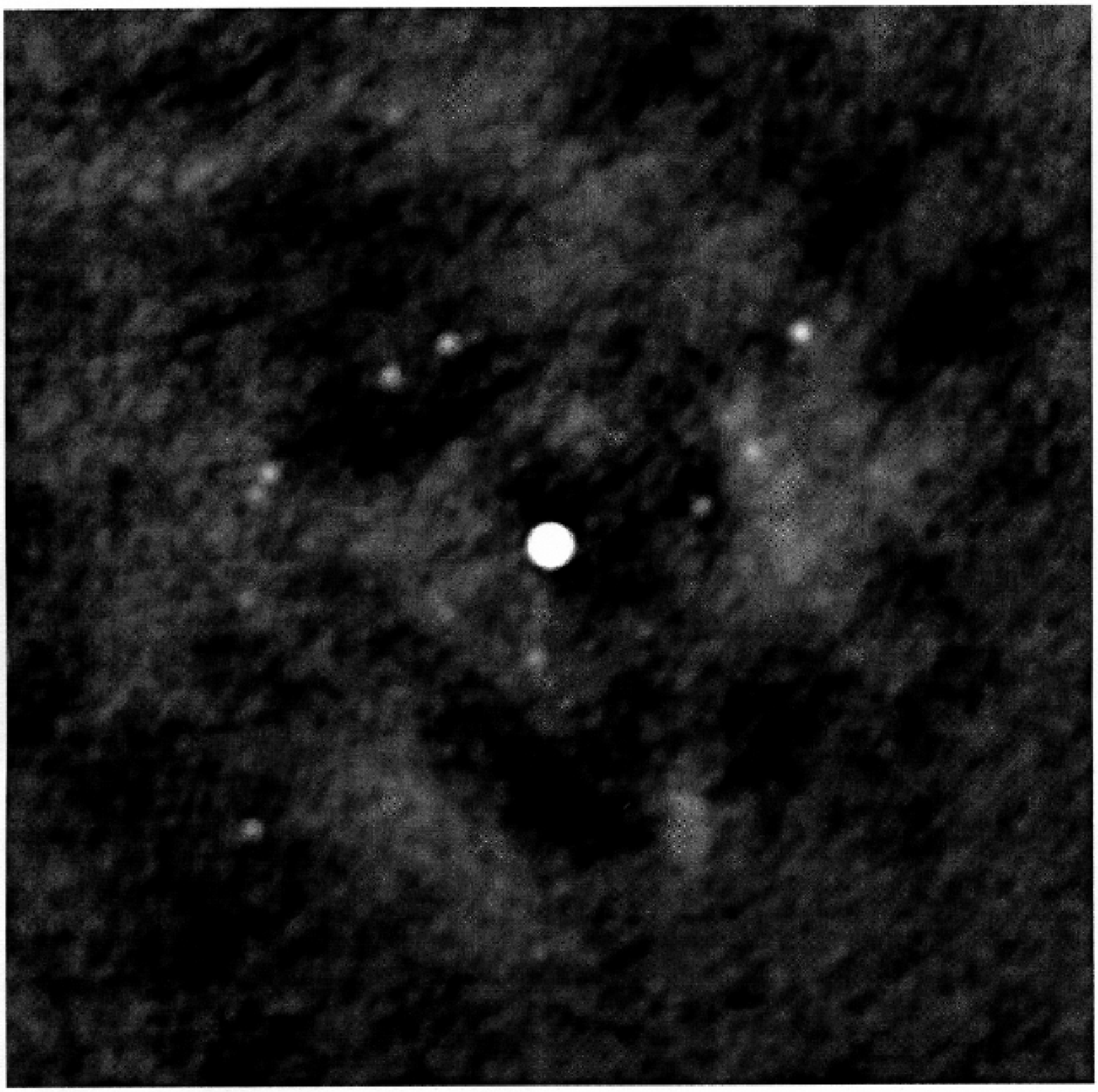
VLA OBSERVATION

(u,v) COVERAGE



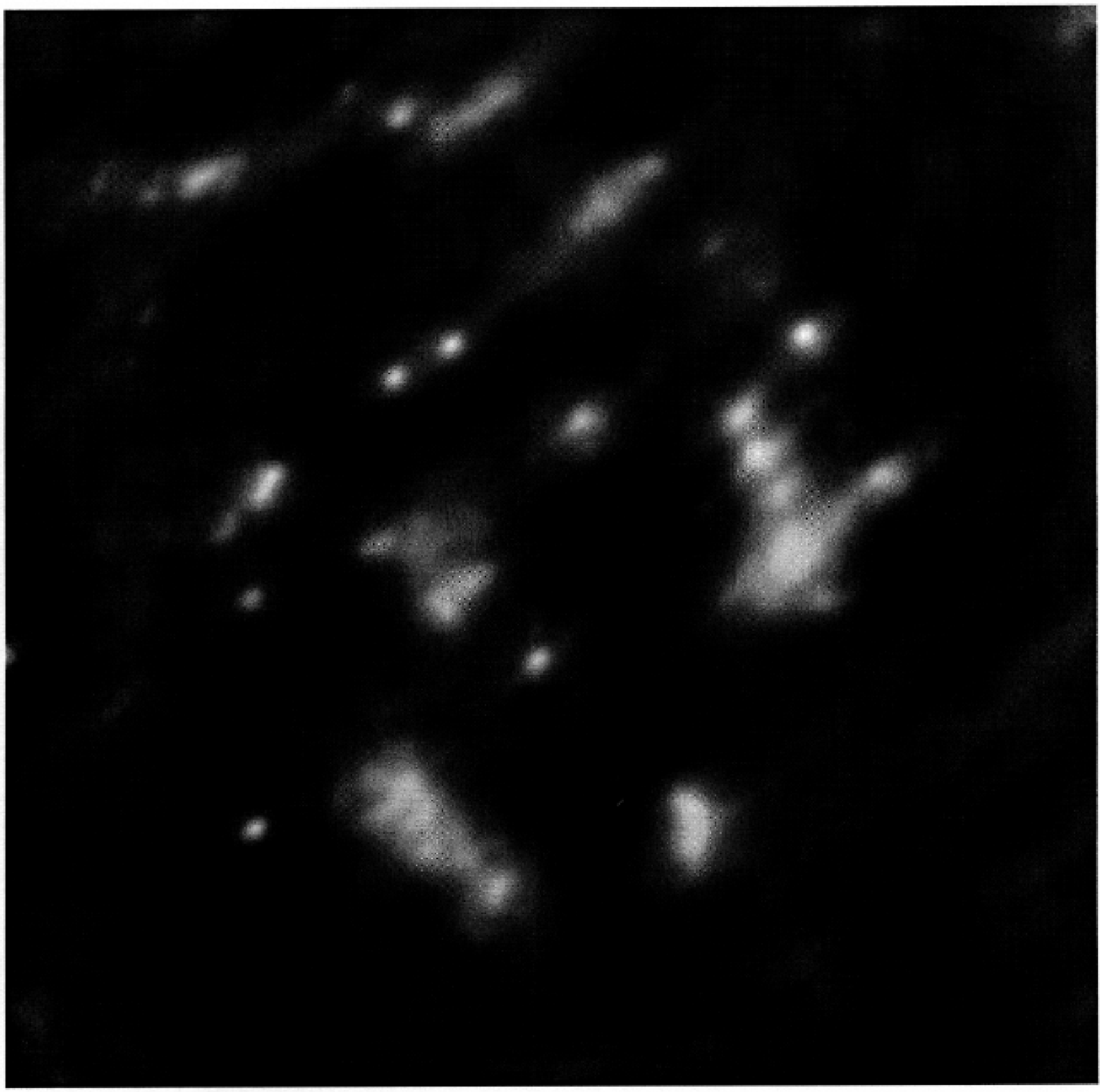
LSI 61° 303

CLEAN RECONSTRUCTION USING AIPS



LSI 61° 303

MAXIMUM LIKELIHOOD RECONSTRUCTION
(100 ITERATIONS)



EXAMPLE FROM RADIO ASTRONOMY

- VLA OBSERVATION OF LSI 61° 303 WITH GOOD (u,v) COVERAGE
- CLEAN AND MAXIMUM LIKELIHOOD SOLUTIONS ARE DIFFERENT

CONCLUSIONS

- INTERFEROMETERS ALREADY EXIST
- SUPERB INTERFEROMETERS FOR OPTICAL / IR ARE UNDER CONSTRUCTION OR AVANCED DESIGN
- THERE IS NO CONSENSUS ABOUT HOW TO HANDLE THE DATA PARTICULARLY FOR IMAGING.
- ALMOST EVERYBODY IN RADIO USE CLEAN
- ALMOST NOBODY IN OPTICAL/IR USE CLEAN (THEY USE R-L / MEM / BAYESIAN)
- SHOULD WE USE CLEAN, MEM, MAXIMUM LIKELIHOOD, BAYESIAN --- ????
- DIFFERENT INSTRUMENTS COULD NEED DIFFERENT ALGORITHMS
- WE NEED REALISTIC SIMULATED DATA TO WORK EVERYBODY WITH SAME DATA (EX: HST)
- MUCH WORK IS STILL TO BE DONE

Integrating polarization into (optical) aperture synthesis: why and how

Jaap Tinbergen

ASTRON, Dwingeloo

In his lecture to the Summer School, Chris Haniff stressed that, at this stage of development of optical interferometry, the most important thing of all is to keep asking: **WHY?** The present paper addresses the questions: **"Why is instrumentation for optical interferometry so different from that for its radio counterpart?"** and: **"Is this difference essential or accidental?"** OK, Chris?

In the section 'Broadbanding an optical complex correlator', I show that an optical ultra-wide-band complex correlator (in the radio-interferometry sense) may be constructed and that such a correlator requires only 4 discrete detectors (pixels) to record all the light. The instantaneous wavelength range could be as wide as 300 nm to over 1 µm (2 octaves). Contrasting this, the usual fringe-recording complex correlator requires several tens of pixels, thus has a multiplex disadvantage whenever system noise is mainly detector noise (and, whenever the fringes are scanned in the time domain using a single detector, much of the light is not even recorded). In addition, forming fringes is essentially a narrow-band process (<20% ?).
Such an ultra-wide-band correlator might have advantages over the usual optical types whenever the exact value of the equivalent wavelength of observation is not of direct interest or should be selectable during an observation without having to retune the system. Examples of such situations might be:

Science:	+ Whenever getting an observation at all depends on having very wide bandwidth. An example might be AGNs. The very small field that goes with wide bandwidth will have to be accepted. + Short exposures (pulsars, cataclysmic variables; planets?)
Instrumentation:	When STJ detectors come into use, intruments such as imaging spectrometers will be built, for which one will want to use as few pixels per correlator as possible. The scheme proposed is basically capable of imaging (=multi-beaming)
Engineering:	+ Fringe trackers (in acquisition mode) could work faster and (in all cases) could work to fainter limiting magnitudes (there is only the one correlation maximum containing all the photons, and this maximum is not blurred by bandwidth effects) + Reference stars for astrometry could be chosen to fainter limiting magnitudes; this might be critical in many applications + Interferometers that can handle images of considerable size will benefit from the fact that only 4 pixels are needed per complex correlator (for reasons that I did not understand, the Large Binocular Telescope was stated to be capable of a larger field than other interferometers)

The proposed ultra-wide-band correlator is based on a phase-switch, which is an old friend from early radio-interferometry and is used in an optical instrument such as FLUOR. While a phase-switch is usually narrow-band (when implemented as a delay-switch), the polarization-optics implementation I propose is ultra-wide-band. Similarly for the phase-shift required to build a complex correlator from 2 simple scalar correlators (generation of the quadrature signal).

One reference has been added: Hamaker 2000A. That paper is the latest and most matured of a series and as such is the best introduction to the use of matrix methods in full-polarization aperture synthesis.

Why polarization:

- EM radiation is *not* a scalar and scalar treatment is incomplete. In particular, as spatial and spectral resolution increase, polarization of celestial objects becomes more visible and more informative (magnetic fields!).
- The best way to implement some instrumental functions may be by polarization components (e.g. ultra-wide-band quadrature-signal generation and an ultra-wide-band halfwave phase switch).
- In any case, astronomical signals being generally of weak and basically unknown polarization, the orthogonal-polarization components of a signal are only weakly correlated and polarization control of the 2 signals fed to a correlator is essential, even if scalar brightness is all one is interested in (i.e. avoiding *photometric errors* caused by *polarization* effects).

There are 3 basic ways to design and/or implement polarization control:

1. Side-step: decide that all one *wishes* to do is make a scalar intensity map of the sky and therefore that all one *needs* to do is to avoid large or unknown loss of correlation; in this case one only needs to equalise the polarization effects (polarizer/depolarizer action, differential phase shifts) in the 2 channels feeding a particular correlator, without actually controlling the polarization effects in quantitative detail; so one equalises instrumental layout in the 2 channels and hopes for the best.
2. Brute force: separate the signal into orthogonal-polarization components (by 2 dipoles, 2 helical antennas or a 2-beam optical polarizer) as early as possible in the system and implement a (scalar) correlator for each combination one needs[1], using the same pair of orthogonal-polarization forms for both feed telescopes ('homogeneous feed system' in radio terms). The interferometer reduces to 2 parallel scalar systems; for the sake of simplicity all the reduction software (selfcal, notably) can then remain in scalar form, yet most (but not all) of the polarization properties of the image can be deduced[2].
3. Fundamental: Use so-called Jones vector and matrix formalism to describe the 'channels', each transporting 2 orthogonal 'eigenmodes' of the polarized feed optics to the correlator, finally expressing the 'complex correlations' (i.e. also containing cross-correlations of the eigenmodes) in terms of a (Wolf) coherency matrix, transforming that to Stokes-parameter 4-vector for astronomical presentation. In radio aperture synthesis, this approach leads to the insight that the 2 channels (telescopes) of the interferometer need not have identical eigenmodes, in fact that greater knowledge of the sky polarization is obtained when the eigenmode pairs are different in the 2 channels ('heterogeneous feed system', with inherent coupling between orthogonal-polarization sky maps). The approach leads to a more complete self-calibration process with better understanding of what is still to be calibrated by external data. AIPS++ can accommodate the full matrix polarization treatment and this possibility is something practicioners of optical aperture synthesis should not discard lightly.

How to integrate polarization:

Although it is not clear (to me!) that optical aperture synthesis can just copy what has been developed for the radio domain (the concept of 'analytic signal', on which Jones formalism is based, plays a different role in the 2 domains – at least as usually employed – and the complex correlator in the optical domain is implicit rather than located within a particular piece of equipment), it is important to use the radio insights and existing reduction techniques where possible. Familiarity with polarization matrix techniques will be a necessity for those specialising in further development of optical aperture synthesis. The potential of polarization in optical signal processing needs to be investigated seriously; the tools exist and the future is to those under-40s who are willing to invest in matrix methods.

An interesting detail is that the matrix formalisms for polarization computations were invented for the optical domain. Radio-astronomical polarimetry used quasi-scalar methods, while mostly employing the same type of correlators as are used in correlation interferometry. After development of the matrix methods for radio correlation interferometry (see Hamaker 1999, 2000 and references therein), optical use looks set for more progress. The basic principles underlying the matrix methods are mentioned below; for more detail, see for example Tinbergen 1996 and its many references.

[1] For complete polarimetry, one will need all 4 independent complex correlations of the 2 × 2 polarized signals, in order to compute sky maps of the 4 Stokes parameters; such correlations may be measured simultaneously or in succession, depending on scientific and technical circumstances.
[2] Not surprisingly, since the 2 scalar systems are independent, the phase difference between the 2 polarizations must be determined by external means

The analytic signal

Use $\sum A_\nu.e^{2\pi i\nu t} = \sum a_\nu.e^{i(2\pi\nu t+\phi_\nu)}$ rather than $\sum a_\nu.\cos(2\pi\nu t+\phi_\nu)$. In optics, one normally takes the real part to obtain a physical signal. In radio practice, the quadrature signal $\sum a_\nu.\sin(2\pi\nu t+\phi_\nu)$ can be constructed and the *analytic signal* $\sum A_\nu.e^{2\pi i\nu t}$ has physical reality. A_ν is complex.

Jones vector and matrix

For polarization, 2 coherent base signals are needed, of orthogonal polarization (usually 2 crossed linear polarizations). A *Jones vector* is a column vector with 2 analytic-signal elements; it completely (amplitude, phase and polarization) represents any **fully-polarized** signal. To represent the transmission of a fully-polarized signal by optical components, 2 × 2 matrices with complex elements are used, so-called *Jones matrices.* Jones calculus is needed to analyse designs of interferometers that use polarization-dependent components (this basically includes all optical aperture synthesis, whether we like it or not).

Stokes vector and Mueller matrix

To describe **partially-polarized** light, one uses the *Stokes vector,* a 4-real-element column vector; the elements are usually denoted by I, Q, U, V and represent **power** in the various polarization forms; there is no phase information, so the system is not suited to describing an interferometer, but is eminently suitable for describing its output: maps of celestial radiation. To describe the effect (in the **power** sense, no phase) of an optical component, 4 × 4 *Mueller matrices* with real elements are used.

Coherency matrix

The *(Wolf) coherency* ***matrix*** contains the same information as the Stokes ***vector***; they are related by

$$\mathbf{E} = \begin{pmatrix} I+Q & U-iV \\ U+iV & I-Q \end{pmatrix}$$

Complex correlation in the basic interferometer

The 4 complex correlations between the 2 × 2 single-polarization analytic signals from a pair of telescopes are the elements of the coherency matrix for that interferometer.

Hamaker, with colleagues Bregman and Sault, represents transmission of signals from feed antennas to the correlator assembly by Jones matrices, the correlator output by coherency matrices. These are then transformed to (complex, hermitian) *Stokes vector visibilities in the u-v plane,* which finally are Fourier-transformed to sky maps of the (real) Stokes parameters and derived quantities. This makes sense: up until the correlator, (so-called 'absolute') phase information is important, afterwards it is irrelevant (only 'differential' phase implicit in the polarization, remains). The treatment allows data-massaging operations such as *self-calibration* to include polarization details and for this reason must be studied by both astronomers and engineers in the (optical) interferometry business (Hamaker 1999, 2000, 2000A and references).

Broadbanding an optical complex correlator

In a radio 'complex correlator', there are 2 tricks which have no direct optical equivalent (none that I have found in the literature, anyway). Optically, one bypasses the need for the tricks when one generates a fringe pattern, taking the amplitude and spatial phase, respectively, of the fringe pattern as the modulus and phase of the complex correlation [the phase becomes of interest when one operates 3 concurrent interferometers in a phase-closure arrangement, a prerequisite for proper (VLBI) imaging].

The 2 tricks are:

1. Multiplication (the essential part of scalar correlation); in radio, this is performed at IF, digitally (1-bit, usually). Optically: by forming $(A+B)^2-(A-B)^2$, either by a 'phase switch' and synchronous demodulation, or implicit in spatial fringing on an array detector.
2. Phase shift of 90°, required to generate the quadrature signal. In radio, implemented by a second mixer, driven by a local-oscillator signal in quadrature. Optically (if used at all): by a path difference.

In contrast to the radio equivalent, the optical arrangements are basically narrowband, hence the spectral dispersion as universal last element in a 'broadband' optical interferometer.

A broad-band 90° phase shift does exist in optics, but it is a *differential* phase shift, in polarization optics; one such component is known as the Fresnel rhomb, which works via the phase jump suffered at total internal reflection. The phase jump depends mainly on refractive index and, since this is a slow function of wavelength in many cases, the 90° phase shift is nearly constant over more than an octave.

In optical astronomy outside aperture synthesis, the choice of bandwidth to use for an observation has often been a delicate balancing act between signal strength and noise of the detector system. One may ask: are there situations in which this holds for optical interferometry and aperture synthesis? The answer might be yes, e.g. when there is no suitable *bright* point source in the isoplanatic patch but there are *faint* ones and there is no laser-star facility; one might then wish to work with, say, 1 octave bandwidth in order to have enough signal within a very short exposure; imaging polarimetry at very high precision and specialized work on time-varying sources (pulsars, cataclysmic variables and planets) might be the most likely clients in astronomical science; in addition, there may be applications in astronomical engineering. The question arises: can the above-mentioned complex-correlator features be generated optically in a wide-band form?

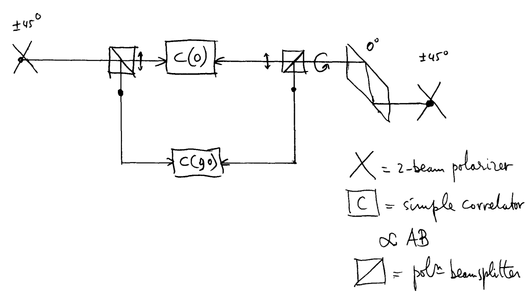


Figure 1: A wide-band optical complex correlator (schematic)

Polarization components might be the answer, since a wide-band 90° phase shifter exists in polarization, while the +/− action is the same as a 0°/180° phase shift and is thus also available as an instantaneous wide-band polarization version (in modulator form also, if required; i.e. a phase switch).

What would such a 'hardwired' complex correlator look like? I have not seen anything resembling this in the (WWW-accessible) literature, so there may be a snag that others have discovered and I have overlooked. Let's do the mental exercise, anyway.

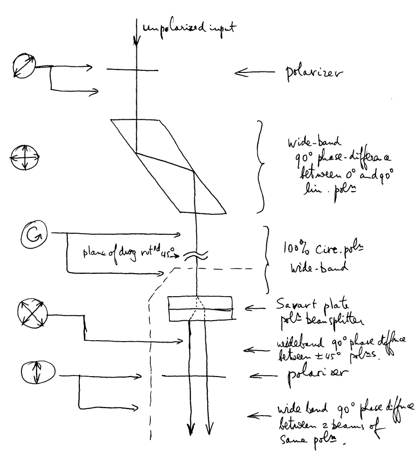


Figure 2: Wide-band optical quadrature-signal generation

A complex correlator consists of 2 scalar correlators, in which one of the signals is correlated both with the direct and the quadrature version of the second signal. An optical wide-band version is shown schematically in Figure 1 and the quadrature signal generation in more detail in Figure 2 . The all-important simple (scalar) correlator is represented here by building blocks (C(0), C(90)); its feasibility is assumed, but is that a reasonable assumption? Figure 3 is an attempt to answer that question. I have no first-hand experience of interferometers (and certainly not of polarization interferometers such as this one), so I am not entirely convinced that the arrangement will work, but if it does, the impact on some areas of optical interferometry could be worthwhile. Figure 3 is schematic only, a realistic device would preferably cater for the complete optical signal (2 polarizations) rather than for the 1 input polarization shown here; one would need the Jones matrix calculus to analyse any configuration proposed.

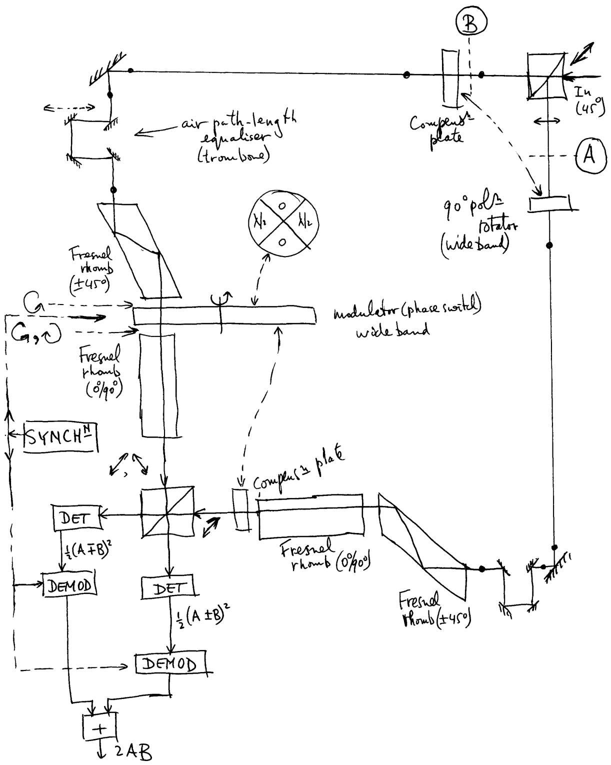
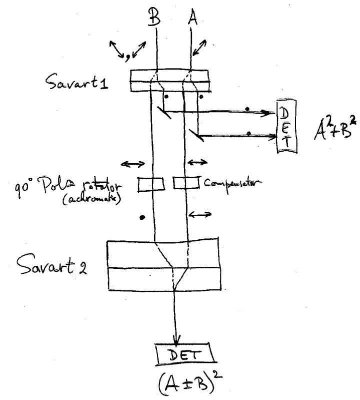


Figure 3: A sketch of a hypothetical wide-band optical (analogue) multiplier

An input signal (shown here as polarized at 45°) is split into 2 coherent components (A and B, to be understood as real signals here, sums of cosines) polarized at 0° and 90°. In aperture synthesis these 2 signals would come from 2 different telescopes and might have the same or (as here) orthogonal polarizations (or be non-orthogonal, in a heterogeneous feed system: e.g. linear from 1 telescope, circular from the other). I have attempted to set up an interferometer that has identical arms in terms of path lengths of air, glass and crystals, and identical phase shifts by reflection at the folding mirrors, yet provides an ultra-wide-band phase switch: the super-achromatic halfwave retarder sectors of the rotating modulator disc. The phase switch works in circular polarization; both it and the circular state of polarization at that point are ultra-wide-band (about 300-1000 nm in a fused-silica, quartz and MgF2 implementation). The AC component of the detector signals is proportional to the product of the input signals and this holds for the entire ultra-wide band, since the paths in the 2 arms are identical in all wavelength-sensitive respects (white-fringe condition, also assumed to hold from the telescopes to the points of definition of the signals A and B). The 4-mirror delay lines are for equalising the air paths; both arms have to have such a system to equalise the phase jumps at the reflections; equalisation by double wedges for both glass and crystals might be necessary, too. The phase switch will have to contain the same total thickness of quartz and MgF_2 in the zero sectors as in the halfwave sectors, but rearranged to give zero retardation for both materials; the compensating plate in the reference arm would have to have the same construction as the zero sectors of the phase switch (the issue of tolerances is swept under the carpet here).

The component I am most worried about is the beam combiner. A calcite version would be best, if feasible (Figure 4). Commercial polarization beamsplitter cubes are never ultra-wide-band, so cannot be used.

What would be the achievable gain over the "widest-narrow-band" systems? Polarization systems always split the total available signal over at least 2, more probably 4 detectors, but the bandwidths can be much wider than in narrow-band systems (an octave or more vs. the widest that would still allow fringe detection) . Since the IR and visible domains have different detector regimes, preferences may differ in the two domains; much will depend on practical details that are entirely beyond this presentation. Since there is no need to record spatial fringes and a single wavelength channel is used, an array detector is not needed. Can one obtain a multiplex advantage by simultaneous multi-beaming (imaging) with an array detector? I have no idea, but the question is important.



A calcite beam combiner?

I suspect that, in a full-polarization version of the multiplier, the other detector (when fed through a second Savart 2), could yield $(A \mp B)^2$

NB. The polarization rotator and compensator are not really needed. Why not?

Figure 4: A possible ultra-wide-band polarization beam combiner?

Conclusions

1. Matrix representation of polarization will be necessary in future aperture synthesis systems, both for data reduction and for the design of instrumentation. Watch radio-astronomy for developments. Both ray-tracing and data-reduction software already allow matrix polarization analysis.
2. I suggest that construction of an ultra-wide-band optical complex correlator is not impossible *a priori* and may be of interest. Matrix methods will be necessary to verify this suggestion.

References

J.P.Hamaker 1999 *Coherency-matrix formulation of self-calibration and some of its salient properties* New Astronomy Reviews **43** 613-616

J.P.Hamaker 2000 *Understanding radio polarimetry IV: the full-coherency analogue of scalar self-calibration: self-alignment, dynamic range and polarimetric fidelity* Astron. Astrophys. Suppl. **143** 515-534

J.P.Hamaker 2000A *Self-calibration of Arrays whose Elements are Strongly Polarized* 'Astronomical Telescopes and Instrumentation 2000', SPIE Conference, Munich (Germany), 25-31 March 2000, in press.

J.Tinbergen 1996 *Astronomical Polarimetry* CUP ISBN 0 521 47531 7

Viewgraphs used at the NEVEC School presentation:

Good introduction
to pol^{n} matrices

Ref.:
J.P. Hamaker 2000A
Self-calibration of Arrays whose Elements are Strongly polarized
SPIE Munich Conf. March 2000

* Polarization ⟺ Somnolescence
* "Radio interferometrists'---- worth stealing"
* SCHOOL
  - Teach (?): polarization
  - Learnt: WHY (why NOT)?

---

* NOT use radio-type correlator?
* NO broadband optical interferometer?
  ("imagers": bandwidth optimization)

WHY?
* NO full-pol^{n} interferometers?
  (not even in Glindemann's > 2006 list)
* designs NOT converging (Haniff)?
* NOT heterodyne?
  (10^{-3} in dig. correl^{n} — Cotton)

RELEVANT to my talk:
+ Different way of combining beams? Eh... yes.
+ Polarization: match or split?
+ Broadbanding: ALCATEL/ESA
+ Multi-beaming / Extended objects ? (cf. Solar speckle) (many isoplanatic patches)
  [speckle ⟷ multi-conjugate AO]
+ Visible wavelengths + 6-mirror Coudé train?
  (from Nasmyth to delay line)
  → fibers → pol^{n} split → $\boxed{\text{radio-type correl}^{n}}$

Conclusions

1) Go Matrix, young (wo)man!
Polarization

2) Steal radio interferometry

3) Ultra-wide-band
(visual = CCD)
- AGN? Other science?
- STJ imaging spectrometers
- Fringetrackers (acqu^{n} mode)?
  Reference stars? Field size (LBT)?
- AOB?

This page was intentionally left blank