RIME: the Radio Interferometer Measurement Equation

Fundamentals of Radio Interferometry

Oleg Smirnov
Rhodes University / SKA SA

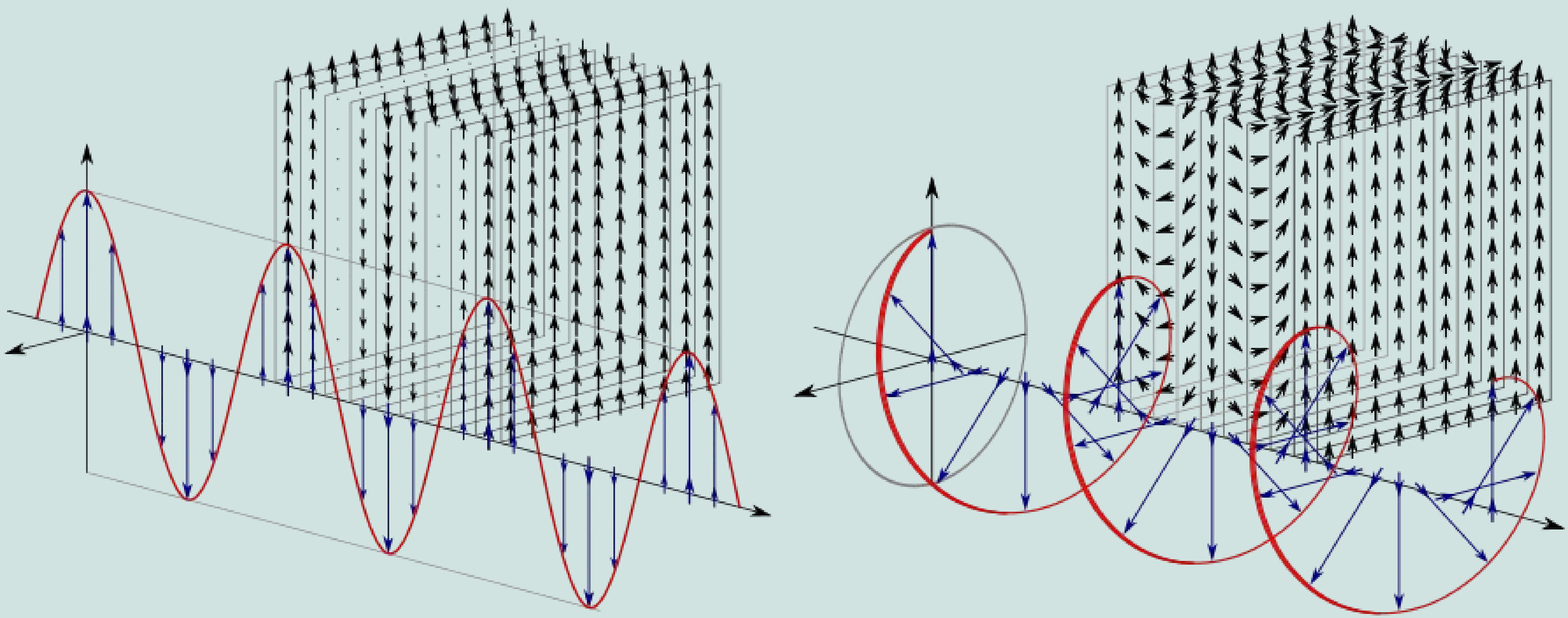
NASSP 2016

The RIME

- You now have some understanding of interferometry
- Let's step back and think about what we're really measuring
- Hence, the radio interferometer measurement equation (RIME)
  - Hamaker, Bregman, Sault 1996
  - Hamaker 2000
- Proper mathematical description of interferometry
  - What we measure, how to calibrate, how to correct

Plane waves

- Electromagnetic plane waves
  - If monochromatic & perfectly polarized:



  - EM field can be described by a complex 2-vector: $$\mathbf{e} = \left[\begin{array}{c} e_x \\ e_y \end{array} \right]$$
  - No *z* component, and same across entire plane

Incoherent radiation

- Radiation from astrophysical sources is neither monochromatic nor perfectly polarized

- Noise! Think of ***e*** waving around randomly
  - But still no *z* component
  - Still the same across the entire plane (a.k.a. it is a transverse EM field)

- Intensity & polarization can then be defined in a statistical sense

Stokes Parameters I

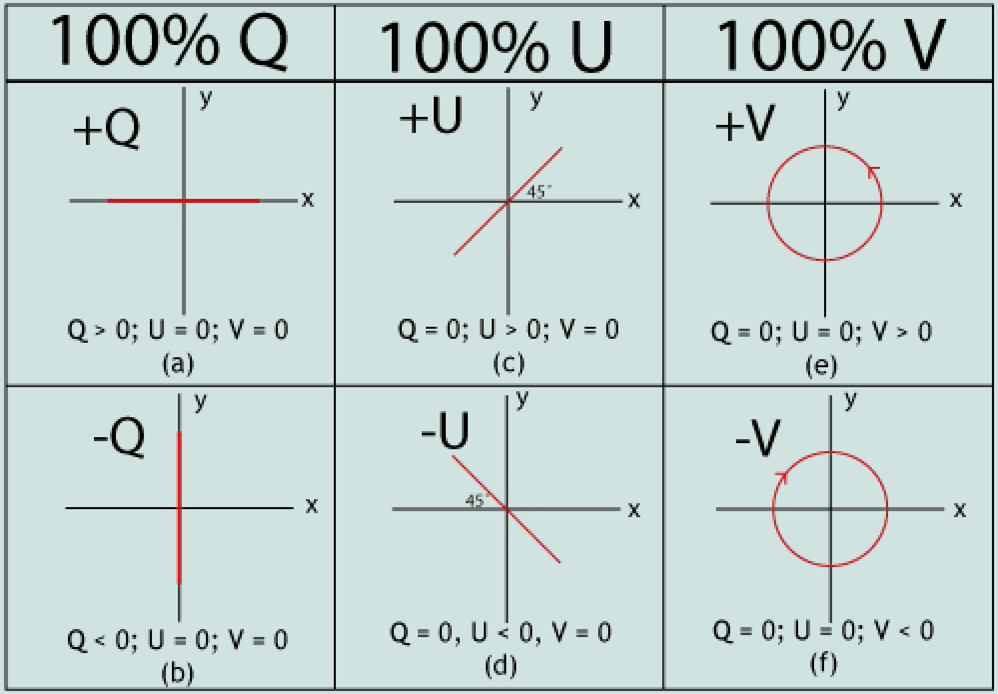
- Stokes parameters are defined in terms of the *coherences*:

$$
\begin{aligned}
I &= \langle e_x e_x^* \rangle + \langle e_y e_y^* \rangle \\
Q &= \langle e_x e_x^* \rangle - \langle e_y e_y^* \rangle \\
U &= \langle e_x e_y^* \rangle + \langle e_y e_x^* \rangle = 2\Re \langle e_x e_y^* \rangle \\
V &= -\imath(\langle e_x e_y^* \rangle - \langle e_y e_x^* \rangle) = 2\Im \langle e_x e_y^* \rangle
\end{aligned}
$$

- The angle brackets operator is an average over a frequency and time bin
- You can also think of e_x and e_y being random variables, and angle brackets denoting *expectation*

Stokes Parameters II

- Perfectly polarized signals:

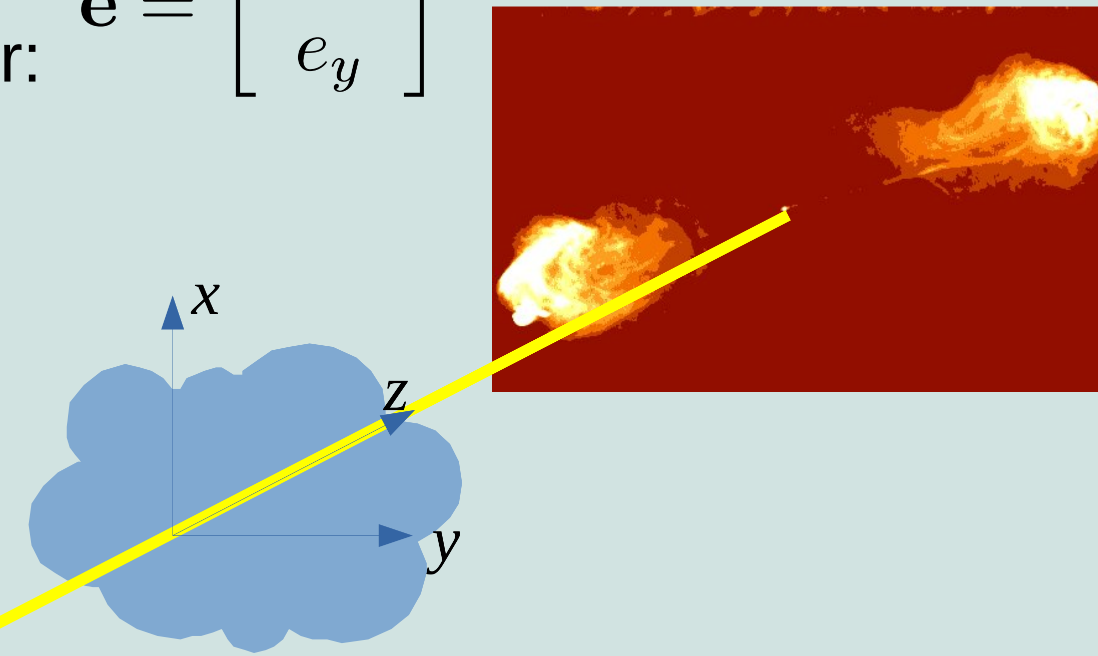


- Partially polarized: ***e*** “waves” in one direction slightly more than the other

Jones Calculus

A transverse EM field can be described by a complex vector: $\mathbf{e} = \begin{bmatrix} e_x \\ e_y \end{bmatrix}$

As the EM wave propagates, the vector changes.



$$\mathbf{e}' = \begin{bmatrix} e'_x \\ e'_y \end{bmatrix}$$

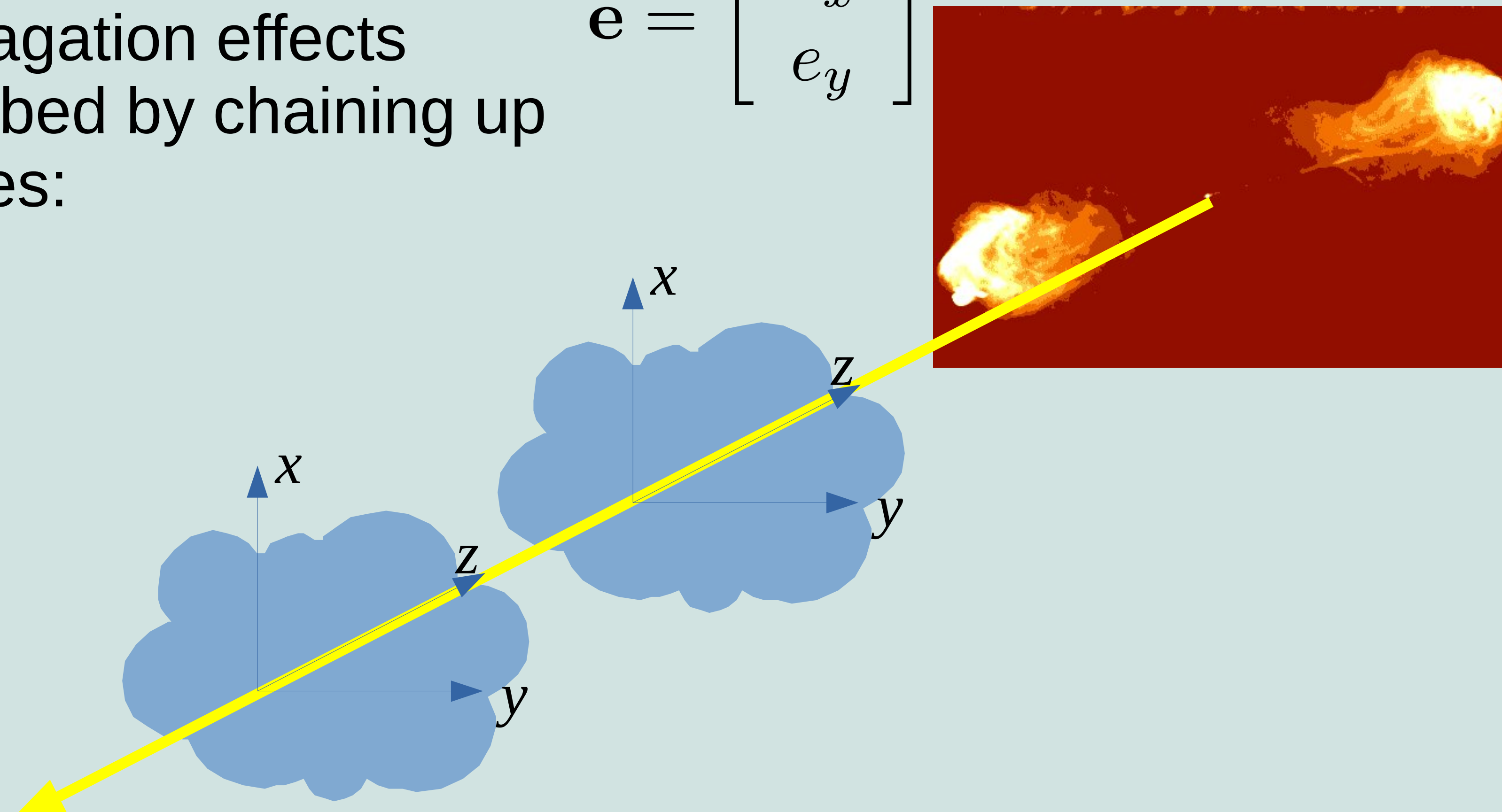
We assume all propagation effects are ***linear***. Any linear transform of a 2-vector can be described by a 2x2 matrix:

$$\mathbf{e}' = \mathbf{J}\mathbf{e} = \begin{bmatrix} j_{11} & j_{12} \\ j_{21} & j_{22} \end{bmatrix} \begin{bmatrix} e_x \\ e_y \end{bmatrix}$$

Jones Chains

Multiple propagation effects can be described by chaining up Jones matrices:

$$\mathbf{e} = \begin{bmatrix} e_x \\ e_y \end{bmatrix}$$

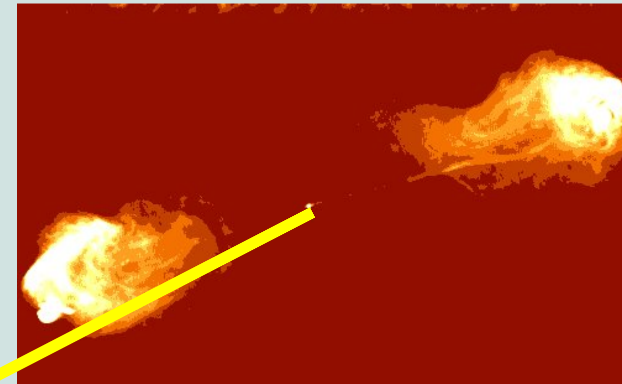


$$\mathbf{e}' = \begin{bmatrix} e'_x \\ e'_y \end{bmatrix}$$

$$\mathbf{e}' = \mathbf{J}_2 \mathbf{J}_1 e$$

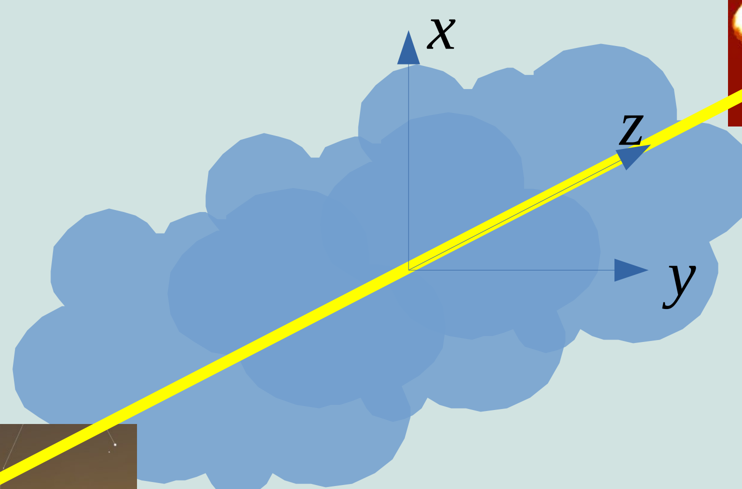
Enter The Antenna

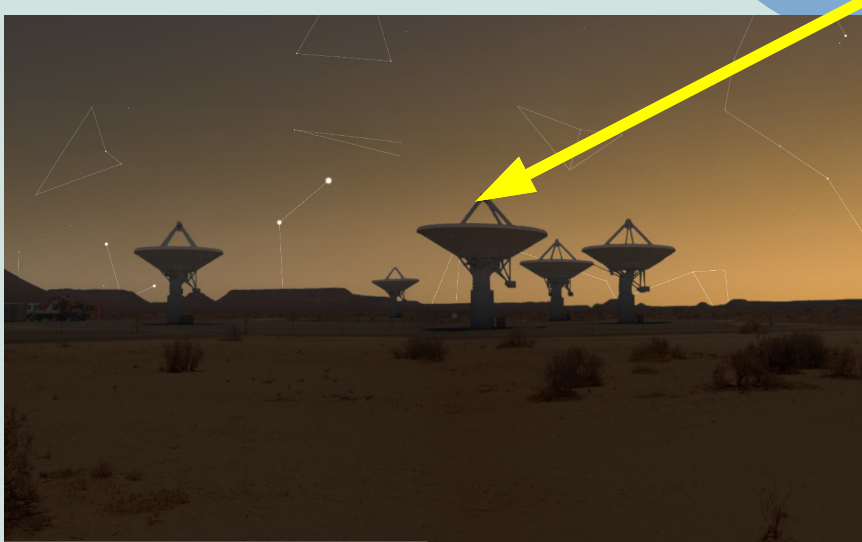
$$\mathbf{e} = \begin{bmatrix} e_x \\ e_y \end{bmatrix}$$



A dual-receptor feed measures two complex voltages (polarizations):

$$\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$$





We may further assume the voltage conversion process is also linear. Therefore we have:

$$\mathbf{v} = \mathbf{J}_n \mathbf{J}_{n-1} ... \mathbf{J}_1 \mathbf{e} = \mathbf{J}\mathbf{e}$$

Jones Zoo

- K-Jones (propagation through free space)

$$\mathbf{K} = K = \left[\begin{array}{cc} \mathrm{e}^{-2\pi\imath\tau/\lambda} & 0 \\ 0 & \mathrm{e}^{-2\pi\imath\tau/\lambda} \end{array} \right] = \mathrm{e}^{-2\pi\imath\tau/\lambda}$$

- P-Jones (parallactic angle rotation)

$$\mathbf{P} = \left[\begin{array}{cc} \cos\gamma & -\sin\gamma \\ \sin\gamma & \cos\gamma \end{array} \right] = \mathrm{Rot}\,\gamma$$

- Zeta-Jones (ionospheric phase delay) $Z = \mathrm{e}^{-\imath\kappa \mathrm{TEC}/\nu}$
- F-Jones (Faraday rotation):

$$\mathbf{F} = \left[\begin{array}{cc} \cos(\mathrm{RM}/\nu^2) & -\sin(\mathrm{RM}/\nu^2) \\ \sin(\mathrm{RM}/\nu^2) & \cos(\mathrm{RM}/\nu^2) \end{array} \right] = \mathrm{Rot}(\mathrm{RM}/\nu^2)$$

Jones Zoo II

- G-Jones (receiver gains)

$$\mathbf{G} = \begin{bmatrix} g_x & 0 \\ 0 & g_y \end{bmatrix}$$

- ...can often be split into a time-variable gain, and a frequency-variable *bandpass*:

$$\mathbf{G} = \begin{bmatrix} g_x(t) & 0 \\ 0 & g_y(t) \end{bmatrix} \qquad \mathbf{B} = \begin{bmatrix} b_x(\nu) & 0 \\ 0 & b_y(\nu) \end{bmatrix}$$

- E-Jones (antenna primary beam)
- D-Jones (polarization leakage)

$$\mathbf{D} = \begin{bmatrix} 1 & d \\ -d & 1 \end{bmatrix}, \quad d \ll 1$$

Jones Sequences

$$\mathbf{v} = \mathbf{J}_n \mathbf{J}_{n-1} ... \mathbf{J}_1 \mathbf{e} = \mathbf{J}\mathbf{e}$$
$$\mathbf{J} = \mathbf{G}\,\mathbf{B}\,\mathbf{D}\,\mathbf{E}\,\mathbf{K}\,\mathbf{P}\,\mathbf{Z}\,\mathbf{F}$$

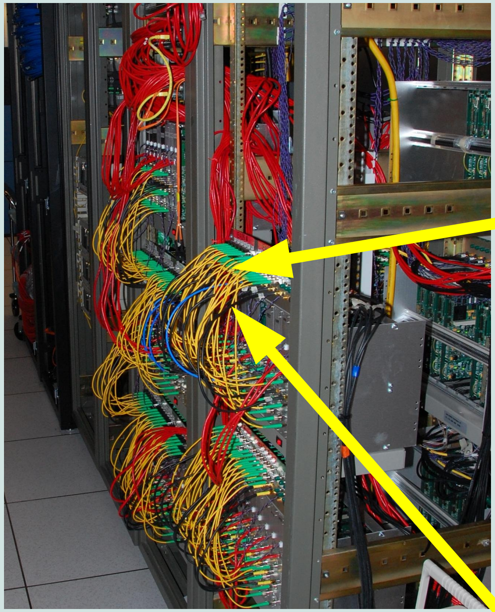
- This is just an example!
- Order is important: matrices don't (in general) commute
  - Must follow physical order of propagation effects
- Some specific matrices do commute
  - Scalar matrix (K-Jones) commutes with everything
  - Diagonal matrices commute among themselves
  - Rotation matrices commute among themselves

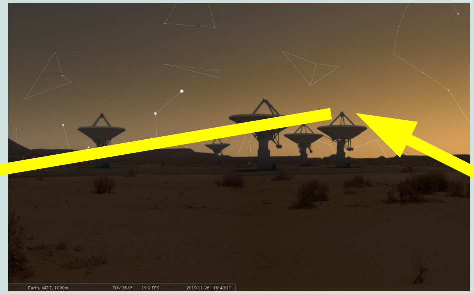
Why Jones?

$$\mathbf{J} = \mathbf{G}\,\mathbf{B}\,\mathbf{D}\,\mathbf{E}\,\mathbf{K}\,\mathbf{P}\,\mathbf{Z}\,\mathbf{F}$$

- To reconstruct the sky, we need to correct for all propagation effects
- Jones calculus allows us to construct a full description of the propagation path
  - some Jones terms are perfectly known *a priori*
  - others can be solved for (i.e. *calibration*)
- Linear algebra tells us how to optimize

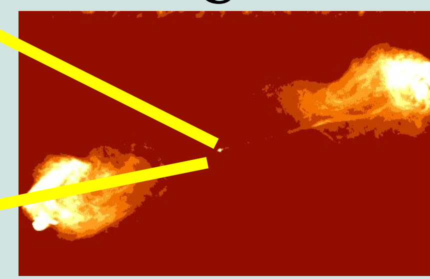
Correlation

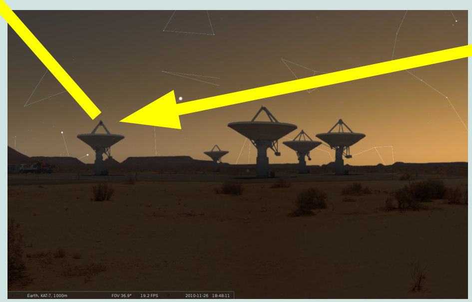




$$\mathbf{v}_p = \mathbf{J}_p \mathbf{e}$$

$$\mathbf{e}$$





$$\mathbf{v}_q = \mathbf{J}_q \mathbf{e}$$

$$V_{11} = \langle v_{p1} v_{q1}^* \rangle$$
$$V_{22} = \langle v_{p2} v_{q2}^* \rangle$$
$$V_{12} = \langle v_{p1} v_{q2}^* \rangle$$
$$V_{21} = \langle v_{p2} v_{q1}^* \rangle$$

A correlator computes four complex pairwise products called *correlations*.

The 2x2 Correlation Matrix

- It proves convenient to arrange the four correlations into a 2x2 ***correlation*** (a.k.a. ***visibility***) ***matrix***, because this can be expressed as a matrix product:

$$\mathbf{V}_{pq} = 2\left[\begin{array}{cc} V_{11} & V_{12} \\ V_{21} & V_{22} \end{array}\right] = 2\left\langle\left[\begin{array}{c} v_{p1} \\ v_{p2} \end{array}\right]\left[\begin{array}{cc} v_{q1}^{*} & v_{q2}^{*} \end{array}\right]\right\rangle = 2\langle\mathbf{v}_p\mathbf{v}_q^H\rangle$$

Enter Jones

- Recalling that

$$\mathbf{v}_p = \mathbf{J}_p \mathbf{e}, \quad \mathbf{v}_q = \mathbf{J}_q \mathbf{e}$$

- ...we have

$$\mathbf{V}_{pq} = 2\langle (\mathbf{J}_p \mathbf{e})(\mathbf{J}_p \mathbf{e})^H \rangle = 2\langle \mathbf{J}_p \mathbf{e}\mathbf{e}^H \mathbf{J}_q^H \rangle = \mathbf{J}_p \, 2\langle \mathbf{e}\mathbf{e}^H \rangle \, \mathbf{J}_q^H$$

$$(\text{using} (\mathbf{AB})^H = \mathbf{B}^H \mathbf{A}^H)$$

- (assuming the Jones matrices are constant over the averaging interval)

The Brightness Matrix

- The inner quantity is called the brightness matrix, and it can be related to the Stokes parameters as:

$$\mathbf{B} = 2\langle \mathbf{e}\mathbf{e}^H \rangle = 2 \begin{bmatrix} e_x e_x^* & e_x e_y^* \\ e_y e_x^* & e_y e_y^* \end{bmatrix} = \begin{bmatrix} I+Q & U+iV \\ U-iV & I-Q \end{bmatrix}$$

- (assuming the Jones matrices are constant over the averaging interval)

The Basic RIME

- This gives us the basic form of the RIME:

Propagation to antenna *p*

Propagation to antenna *q*

$$\mathbf{V}_{pq} = \mathbf{J}_p \mathbf{B} \mathbf{J}_q^H$$

Measured correlation

Intrinsic source parameters

The Onion RIME

- Recall Jones chains:

$$\mathbf{J}_p = \mathbf{J}_{p,n}\mathbf{J}_{p,n-1}\ldots\mathbf{J}_{p,1},$$

- This gives us the “onion form” of the RIME

$$\mathbf{V}_{pq} = \mathbf{J}_{p,n}(\ldots(\mathbf{J}_{p,1}\mathbf{B}\mathbf{J}_{q,1}^H)\ldots)\mathbf{J}_{q,n}^H$$

Specific vs. General RIMEs

- The above is “**<u>The</u>** RIME”, i.e. a general formalism
- In practice, we string together some specific Jones terms (depending on how accurately we need to model the an observation), and call it **<u>a</u>** RIME, or **<u>a</u>** measurement equation
- E.g. a simple RIME for a single point source, with receiver gains:

$$\mathbf{V}_{pq} = \mathbf{G}_p K_p \mathbf{B} K_q^H \mathbf{G}_q^H$$

- Or an even simpler RIME with an ideal instrument:

$$\mathbf{V}_{pq} = K_p \mathbf{B} K_q^H$$

(recall that K-Jones is propagation through free space, a.k.a geometric delay)

Geometric Delay

$$K_p = \mathrm{e}^{-2\pi i(\mathbf{u}_p \cdot \sigma)/\lambda}$$

$$\mathbf{V}_{pq} = K_p \mathbf{B} K_q^H = \mathbf{B} K_p K_q^H$$

$$= \mathbf{B}\,\mathrm{e}^{-2\pi i((\mathbf{u}_p - \mathbf{u_q})\cdot\sigma)/\lambda}$$

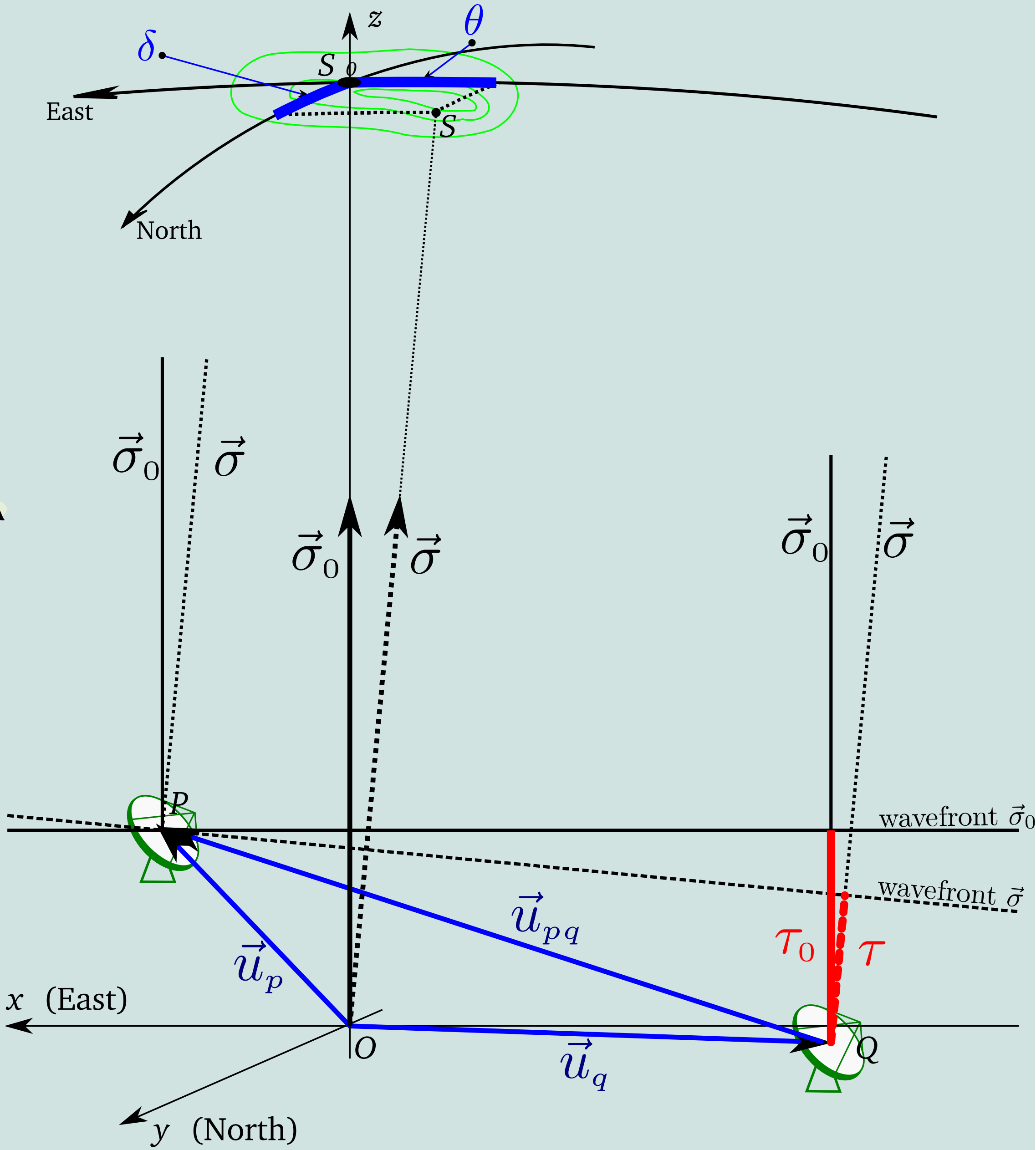
$$= \mathbf{B}\,\mathrm{e}^{-2\pi i(u_{pq}l + v_{pq}m + w_{pq}n)/\lambda}$$

- with the unit direction vector

$$\sigma = (l, m, n)^T$$

- *l,m,n* are also called the *direction cosines*

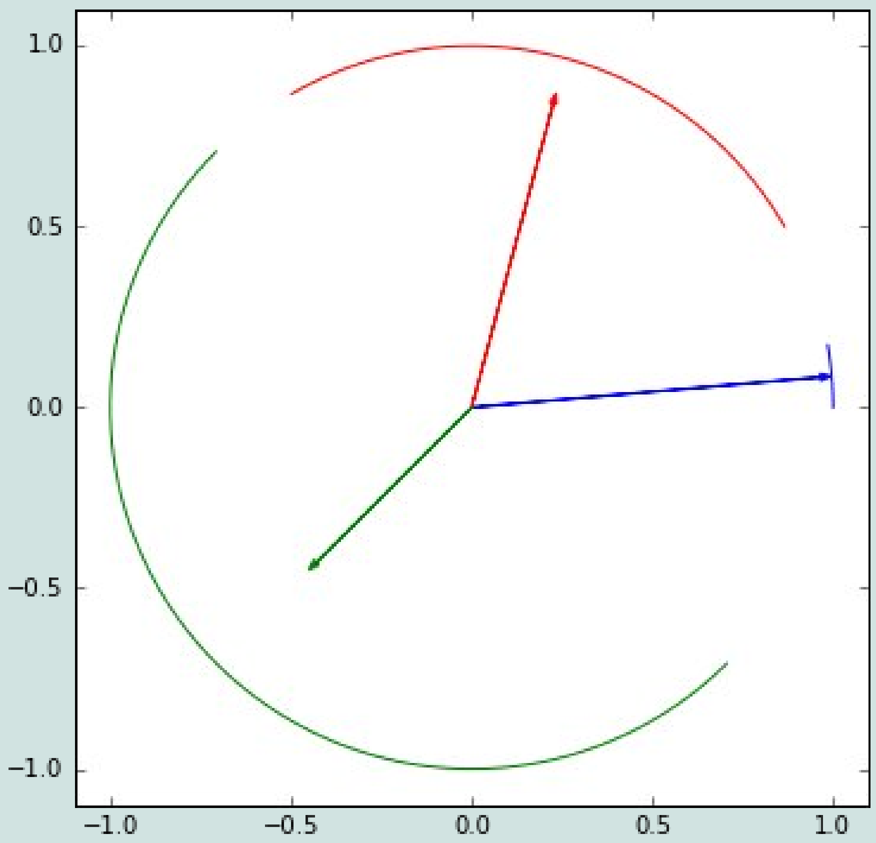
$$(n = \sqrt{1 - l^2 - m^2})$$



Averaging & Smearing

- Our derivation of the RIME assumed a constant Jones term over the averaging interval
- This is not always safe, e.g. *K*-Jones varies (rotates) with frequency and time:

$$K_p = \mathrm{e}^{-2\pi i(\mathbf{u}_p \cdot \sigma)/\lambda}$$



- The effect of averaging a rotating complex vector is a reduction in amplitude
- This is known as [time/frequency] ***smearing***

Fringe Stopping

- We want to minimize phase rotation in our region of interest i.e. around the direction $\sigma_0 = (0,0,1)^T$

$$K_p K_q^* = \mathrm{e}^{-2\pi i(u_{pq}l + v_{pq}m + w_{pq}n)/\lambda}$$

- We can do this by having the correlator insert an artificial phase delay of $\mathrm{e}^{2\pi i w_{pq}/\lambda}$
- This known as ***fringe stopping***
- For a fringe stopping correlator, the K-Jones term is

$$K_p = \mathrm{e}^{-2\pi i(u_p l + v_p m + w_p(n-1))/\lambda}$$

- This ensures constant phase towards the direction of interest, and only slowly rotating phase near it

The All-Sky RIME

- Until now we have only considered a single point source with brightness matrix **B**
- The real sky is a brightness distribution as a function of direction: $\mathbf{B}(\sigma)$
- The signal path from each direction is, in principle, different: $\mathbf{J}_p(\sigma)$
- Signals from different directions add up linearly, and we have an integral over a sphere:

$$\mathbf{V}_{pq} = \iint\limits_{4\pi} \mathbf{J}_p(\sigma)\mathbf{B}(\sigma)\mathbf{J}_q^H(\sigma)\mathrm{d}\sigma$$

The All-Sky RIME II

Let's recast this as a 2D integration over the *l,m* plane:

$$\mathbf{V}_{pq} = \iint\limits_{lm} \mathbf{J}_p(l,m)\mathbf{B}(l,m)\mathbf{J}_q^H(l,m)\frac{\mathrm{d}\sigma}{\mathrm{d}l\mathrm{d}m}\mathrm{d}l\mathrm{d}m$$

$$= \iint\limits_{lm} \mathbf{J}_p(l,m)\mathbf{B}(l,m)\mathbf{J}_q^H(l,m)\frac{1}{n}\mathrm{d}l\mathrm{d}m,$$

- Each Jones matrix is a chain: $\mathbf{J}_p(l,m) = \mathbf{J}_{p,n}...\mathbf{J}_{p,1}$
- ...which we can split into *direction-dependent* (DD) and *direction-independent* (DI) parts:

$$\mathbf{J}_p(l,m) = (\mathbf{J}_{p,n}...\mathbf{J}_{p,k+1})(\mathbf{J}_{p,k}...\mathbf{J}_{p,2})K_p$$

$$\mathbf{J}_p(l,m) = \mathbf{G}_p\mathbf{E}_p(l,m)K_p(l,m)$$

The All-Sky RIME: a 2D Fourier Transform?

- We now have:

$$\mathbf{V}_{pq} = \mathbf{G}_p \left(\iint\limits_{lm} \frac{1}{n} \mathbf{E}_p(l,m) \mathbf{B}(l,m) \mathbf{E}_q^H(l,m) \mathrm{e}^{-2\pi i(u_{pq}l + v_{pq}m + w_{pq}(n-1))/\lambda} \mathrm{d}l \mathrm{d}m \right) \mathbf{G}_q^H$$

- Let's formally define a W-Jones as

$$W_p = \frac{1}{\sqrt{n}} \mathrm{e}^{w_p(n-1)},$$

- ...and absorb it into E-Jones:

$$\mathbf{V}_{pq} = \mathbf{G}_p \left(\iint\limits_{lm} \mathbf{B}_{pq}(l,m) \mathrm{e}^{-2\pi i(u_{pq}l + v_{pq}m)/\lambda} \mathrm{d}l \mathrm{d}m \right) \mathbf{G}_q^H$$

$$\text{with} \quad \mathbf{B}_{pq} = \mathbf{E}_p \mathbf{B} \mathbf{E}_q^H$$

No DD Effects = Fourier Transform

- In the absence of DD effects, the integral is a true 2D FT:

$$\mathbf{V}_{pq} = \mathbf{G}_p \left(\iint_{lm} \mathbf{B}(l,m) \mathrm{e}^{-2\pi i (u_{pq} l + v_{pq} m)/\lambda} \mathrm{d}l \mathrm{d}m \right) \mathbf{G}_q^H$$

- If we can estimate **G** (see Calibration), we can form corrected visibilities

$$\mathbf{V}_{pq}^{(\mathrm{corr})} = \tilde{\mathbf{G}}_p^{-1} \mathbf{V}_{pq} \tilde{\mathbf{G}}_q^{-H}$$

- ...which correspond to a 2D FT of the underlying sky
- But DD effects always present (e.g. antenna primary beam)

Trivial DD Effects

- Trivial DD effects are those that do not vary with time, an are the same across all antennas:

$$\mathbf{E}_p(t) \equiv \mathbf{E}$$

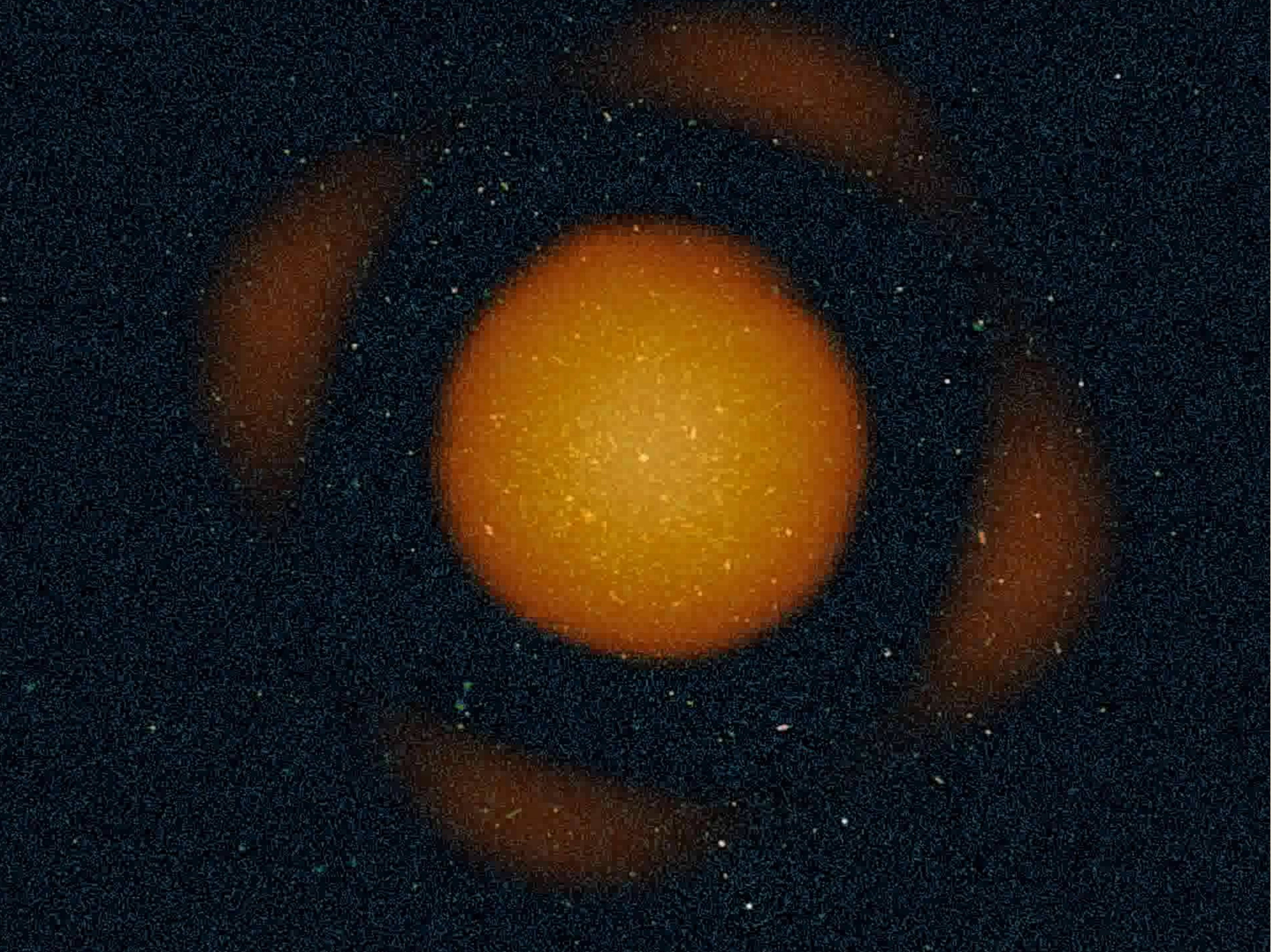
- We can then define an ***apparent sky***:

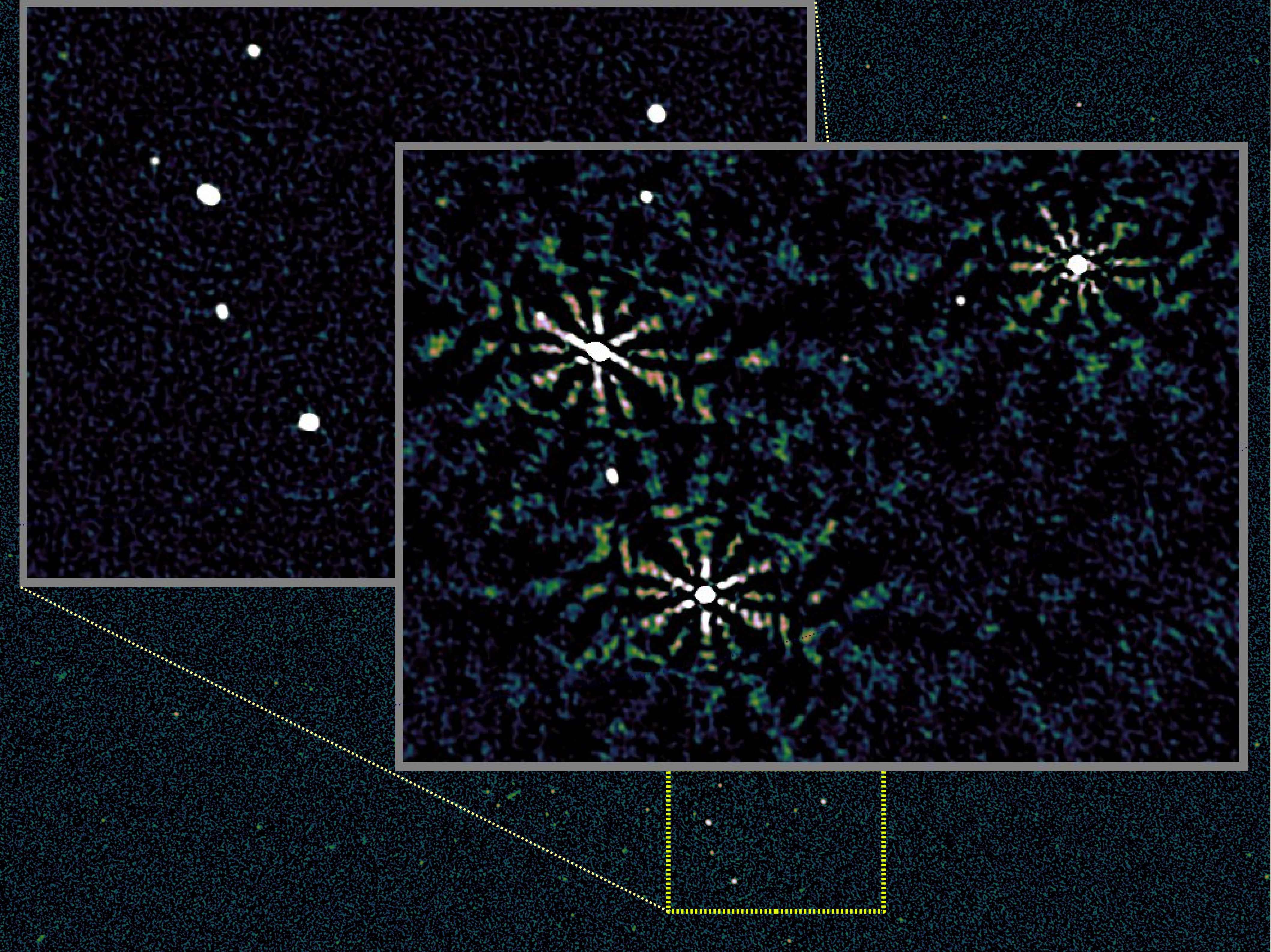
$$\mathbf{B}_{\mathrm{app}} = \mathbf{E}\mathbf{B}\mathbf{E}^H$$

- ...and treat our measurement as true 2D FT of the apparent sky
- This is the "classical regime" of interferometry (aka selfcal, or 2GC): neglect DDEs and recover the apparent sky

Non-trivial DD Effects

- The classical regime breaks down when DDEs are not trivial
- For example, W-Jones is only trivial with either
  - (a) narrow fields of view: $n = \sqrt{1 - l^2 - m^2} \to 1$
  - or (b) coplanar arrays: $w \equiv 0$
- Hence, wide-field problem (Chapter 5)
- Primary beam (E-Jones) is only trivial when
  - Dishes are identical (they're not)
  - Pointing is perfect (it isn't)
  - Sky does not rotate (equatorial or tri-axis mount)





Mueller Calculus

- An alternative to Jones calculus: often found in imaging literature
- Instead of using a brightness matrix, we pack the Stokes parameters into a ***Stokes vector***:

$$\mathbf{b} = \begin{bmatrix} I \\ Q \\ U \\ V \end{bmatrix}$$

- We also define a *measured* Stokes vector for each baseline:

$$\mathbf{b}_{pq} = \begin{bmatrix} I' \\ Q' \\ U' \\ V' \end{bmatrix} = \frac{1}{2} \begin{bmatrix} V_{11} + V_{22} \\ V_{11} - V_{22} \\ V_{12} + V_{21} \\ -\imath(V_{12} - V_{21}) \end{bmatrix}$$

Mueller Matrices

- Since we assumed the system is linear, there must be a 4x4 matrix relating the two vectors:

$$\mathbf{b}_{pq} = \mathbf{M}_{pq}\mathbf{b}$$

- This is called the ***Mueller matrix*** of baseline *pq*
- Useful, because it emphasizes the direct relationship between observed and intrinsic Stokes parameters
- Often found in imaging literature

Deriving Mueller Matrices

- Using the Kronecker (outer) product, we can stack the visibilities into a 4-vector:

$$\mathbf{v}_{pq} = 2 \begin{bmatrix} V_{11} \\ V_{12} \\ V_{21} \\ V_{22} \end{bmatrix} = 2\mathbf{v}_p \otimes \mathbf{v}_q^*$$

- By analogy, we can define a coherency vector

$$\mathbf{x} = 2\mathbf{e} \otimes \mathbf{e}^* = \begin{bmatrix} I + Q \\ U + \imath V \\ U - \imath V \\ I - Q \end{bmatrix}$$

- ...and using the mixed product identity: $(\mathbf{AB}) \otimes (\mathbf{CD}) = (\mathbf{A} \otimes \mathbf{B})(\mathbf{C} \otimes \mathbf{D})$

- we get:

$$\mathbf{v}_{pq} = 2(\mathbf{J}_p\mathbf{e}) \otimes (\mathbf{J}_p\mathbf{e})^* = (\mathbf{J}_p \otimes \mathbf{J}_q^*)(2\mathbf{e} \otimes \mathbf{e}^*) = (\mathbf{J}_p \otimes \mathbf{J}_q^*)\mathbf{x}$$

Deriving Mueller Matrices II

- The coherency vector relates to the Stokes vector via a conversion matrix **S**:

$$\mathbf{x} = \begin{bmatrix} I+Q \\ U+\imath V \\ U-\imath V \\ I-Q \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & \imath \\ 0 & 0 & 1 & -\imath \\ 1 & -1 & 0 & 0 \end{bmatrix} \begin{bmatrix} I \\ Q \\ U \\ V \end{bmatrix} = \mathbf{S}\,\mathbf{b}$$

- The opposite relation is

$$\mathbf{b} = \begin{bmatrix} I \\ Q \\ U \\ V \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & -1 \\ 0 & 1 & 1 & 0 \\ 0 & -\imath & \imath & 0 \end{bmatrix} \begin{bmatrix} I+Q \\ U+\imath V \\ U-\imath V \\ I-Q \end{bmatrix} = \mathbf{S}^{-1}\mathbf{x}$$

Deriving Mueller Matrices III

- We can now write the measured Stokes vector as

$$\mathbf{b}_{pq} = \mathbf{S}^{-1}\mathbf{v}_{pq}$$

- ...and the full RIME becomes:

$$\mathbf{b}_{pq} = \mathbf{S}^{-1}(\mathbf{J}_p \otimes \mathbf{J}_q^*)\mathbf{S}\,\mathbf{b}$$

- And the Mueller matrix of baseline *pq* is given by:

$$\mathbf{M}_{pq} = \mathbf{S}^{-1}(\mathbf{J}_p \otimes \mathbf{J}_q^*)\mathbf{S}$$