ALS calibration: analytic expressions for the antenna gains of a two and three element east-west interferometer.

T. L. Grobler[1,2], O. M. Smirnov*[1,2], C.D. Nunhokee[1]

[1]*Department of Physics and Electronics, Rhodes University, PO Box 94, Grahamstown, 6140, South Africa,*

trienkog@gmail.com,osmirnov@gmail.com,ridhima_24@yahoo.com

[2]*SKA South Africa, 3rd Floor, The Park, Park Road, Pinelands, 7405, South Africa*

Abstract

Studying a simple problem first, for which an analytic solution exists, can greatly improve our understanding of a complex problem. An undesired result of calibrating with an incomplete sky model is calibration systematics. Studying the systematics contained in the antenna gains of a simple east-west interferometer could aid us in understanding the artefacts associated with calibration in general. This paper aims to derive analytic expressions of the antenna gain vectors that were obtained during the calibration (using an incomplete sky model) of a two and three element east-west interferometer.

1 Introduction

The analytic expressions derived in this paper turned out to be one of the fundamental building blocks needed to explain spurious calibration artefacts or "ghost sources" [1]. Since these analytic equations were not recorded in the original paper we thought it worthwhile to present it here. These analytic equations will not only help a reader to better understand the theory presented in [1], but could also turn out to be useful when solving another unrelated problem. The interferometer studied here is assumed to be perfect, i.e. the antenna gains are assumed to be unity. Moreover, the observed sky is assumed to consist of two point sources, one in the phase center and one off-center, while the calibration sky model only contains the central source.

2 Calibration

Unpolarized calibration entails finding the antenna gains $\boldsymbol{g} = (g_1, g_2, \cdots, g_n)^T$ which minimizes [1, 3]:

$$||\boldsymbol{\mathcal{R}} - \boldsymbol{G}\boldsymbol{\mathcal{M}}\boldsymbol{G}^H|| = ||\boldsymbol{\mathcal{R}} - \boldsymbol{g}\boldsymbol{g}^H \odot \boldsymbol{\mathcal{M}}|| = ||\boldsymbol{\mathcal{R}} - \boldsymbol{\mathcal{G}} \odot \boldsymbol{\mathcal{M}}||, \tag{1}$$

where $\boldsymbol{\mathcal{R}}$ is the observed unpolarized visibilities matrix, $\boldsymbol{\mathcal{M}}$ is the calibration model matrix, $\boldsymbol{\mathcal{G}} \odot \boldsymbol{\mathcal{M}}$ is the calibrated visibility matrix, $\boldsymbol{g}$ is the antenna gains matrix, $\boldsymbol{G}$ is equal to diag($\boldsymbol{g}$), $\boldsymbol{\mathcal{G}} = \boldsymbol{G}\boldsymbol{1}\boldsymbol{G}^H = \boldsymbol{g}\boldsymbol{g}^H$, "$H$" denotes the Hermitian transpose and "$\odot$" denotes the Hadamard product. The entries of $\boldsymbol{\mathcal{R}}$, $\boldsymbol{\mathcal{M}}$ and $\boldsymbol{\mathcal{G}}$ are respectively denoted by r_{pq}, m_{pq} and g_{pq}. If the calibration model consists of only one point source located at the phase center (a 1Jy source), then Eq. 1 simplifies and becomes

$$||\boldsymbol{\mathcal{R}} - \boldsymbol{\mathcal{G}}||, \tag{2}$$

since $\boldsymbol{\mathcal{M}} = \boldsymbol{1}$. For the sake of simplicity let us assume that the interferometer estimated the autocorrelations perfectly. The $\boldsymbol{\mathcal{G}}$ that minimizes Eq. 2 can now be calculated by using Alternating Least Squares (ALS) calibration and is thus equal to [2]

$$\boldsymbol{\mathcal{G}} = \lambda \mathbf{x}\mathbf{x}^H, \tag{3}$$

where λ is the largest eigenvalue of $\boldsymbol{\mathcal{R}}$ and $\mathbf{x}$ is its associated (normalized) eigenvector. The antenna gain vector $\boldsymbol{g}$ is therefore equal to

$$\boldsymbol{g} = \sqrt{\lambda}\mathbf{x}. \tag{4}$$

3 Eigenvalues and Eigenvectors

ALS depends on eigenvalue decomposition and as such a review of eigenvalues and eigenvectors is required. The eigenvalue λ, which is a single number, and its associated eigenvector $\mathbf{x}$ of a square matrix $\boldsymbol{A}$ satisfies

$$\boldsymbol{A}\mathbf{x} = \lambda \mathbf{x}.$$

The eigenvalues of a square matrix $\boldsymbol{A}$ is calculated by determining the roots of the characteristic equation

$$|\boldsymbol{A} - \lambda \boldsymbol{I}| = 0.$$

978-1-4673-5225-3/14/$31.00 ©2014 IEEE

There exists closed form expressions for the eigenvalues and eigenvectors of 2×2 and 3×3 matrices. The eigenvalues and eigenvectors of 2×2 and 3×3 matrices are especially important in this paper and therefore discussed in greater detail in Section 3.1 and Section 3.2.

3.1 A 2×2 matrix

Assume that

$$\boldsymbol{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}, \tag{5}$$

and that $a_{21} \neq 0$ then the eigenvalues of $\boldsymbol{A}$ are equal to [4]

$$\lambda_1 = \frac{\mathrm{tr}(\boldsymbol{A})}{2} + \sqrt{\frac{\mathrm{tr}(\boldsymbol{A})^2}{4} - |\boldsymbol{A}|}, \tag{6}$$

$$\lambda_2 = \frac{\mathrm{tr}(\boldsymbol{A})}{2} - \sqrt{\frac{\mathrm{tr}(\boldsymbol{A})^2}{4} - |\boldsymbol{A}|}, \tag{7}$$

where $\mathrm{tr}(\boldsymbol{A}) = a_{11} + a_{22}$ and $|\boldsymbol{A}| = (a_{11}a_{22}) - (a_{12}a_{21})$. The unnormalized eigenvectors of λ_1 and λ_2 (not unique) are respectively equal to [4]

$$\mathbf{x}^{(1)} = [\lambda_1 - a_{22} \; a_{21}]^T, \tag{8}$$

$$\mathbf{x}^{(2)} = [\lambda_2 - a_{22} \; a_{21}]^T. \tag{9}$$

3.2 A 3×3 matrix

Assume that

$$\boldsymbol{A} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}, \tag{10}$$

then the characteristic equation of $\boldsymbol{A}$ is equal to

$$|\boldsymbol{A} - \lambda \boldsymbol{I}| = \lambda^3 - \mathrm{tr}(\boldsymbol{A})\lambda^2 - \frac{1}{2}(\mathrm{tr}(\boldsymbol{A}^2) - \mathrm{tr}^2(\boldsymbol{A}))\lambda - |\boldsymbol{A}|. \tag{11}$$

The eigenvalues $(\lambda_1, \lambda_2, \lambda_3)$ of $\boldsymbol{A}$ are equal to the roots of Eq. 11. An unnormalized eigenvector $\mathbf{x}^{(i)}$ corresponding to λ_i can be calculated with

$$\mathbf{x}^{(i)} = [(\boldsymbol{A}_1 - \lambda_i e_1) \times (\boldsymbol{A}_2 - \lambda_i e_2)]^T, \tag{12}$$

where $\boldsymbol{A}_x$ is the x-th row of $\boldsymbol{A}$, $e_1 = [1 \; 0 \; 0]$, $e_2 = [0 \; 1 \; 0]$ and $\times$ denotes the vector cross product. Eq. 12 is only valid if $\boldsymbol{A}_1 - \lambda_i e_1$ and $\boldsymbol{A}_2 - \lambda_i e_2$ are linearly independent vectors [5].

4 Analytic expressions of the antenna gains

For the sake of simplicity let us assume that the true sky equals $I_{\mathcal{R}}(\boldsymbol{s}) = A_1\delta(\boldsymbol{s}) + A_2\delta(\boldsymbol{s} - \boldsymbol{s}_0)$ (consists out of two point sources) and that the calibrated sky model equals $I_{\mathcal{M}}(\boldsymbol{s}) = A_1$, where A_1 and A_2 are source fluxes and $\boldsymbol{s}_0 = (l_0, m_0)$ is the direction cosine position vector of the off-center source. Note that $A_1 = 1$ for the remainder of the paper.

4.1 Two element interferometer

When the sky is equal to $I_{\mathcal{R}}(\boldsymbol{s})$ then the observed visibility matrix $\boldsymbol{\mathcal{R}}_{pq}(\boldsymbol{b}_{pq})$ is equal to

$$\begin{aligned} \boldsymbol{\mathcal{R}}_{pq}(\boldsymbol{b}_{pq}) &= \begin{bmatrix} V_{pp}^{\mathcal{R}} & V_{pq}^{\mathcal{R}} \\ V_{qp}^{\mathcal{R}} & V_{qq}^{\mathcal{R}} \end{bmatrix} \\ &= \begin{bmatrix} A_1 + A_2 & A_1 + A_2 e^{-2\pi i \boldsymbol{b}_{pq} \cdot \boldsymbol{s}_0} \\ A_1 + A_2 e^{2\pi i \boldsymbol{b}_{pq} \cdot \boldsymbol{s}_0} & A_1 + A_2 \end{bmatrix}. \end{aligned} \tag{13}$$

where $\boldsymbol{b}_{pq} = (u_{pq}, v_{pq})$, and p and q are antenna indexes.

Now let

$$c(u_{pq}, v_{pq}) = A_1 + A_2 e^{2\pi i \boldsymbol{b}_{pq}\cdot \boldsymbol{s}_0} \tag{14}$$
$$= A_1 + A_2 \cos(2\pi \boldsymbol{b}_{pq} \cdot \boldsymbol{s}_0) + i \sin(2\pi \boldsymbol{b}_{pq} \cdot \boldsymbol{s}_0), \tag{15}$$
$$c_0 = A_1 + A_2 e^{2\pi i \boldsymbol{0}\cdot \boldsymbol{s}_0} = A_1 + A_2, \tag{16}$$
$$|c(u_{pq}, v_{pq})| = \sqrt{A_1^2 + A_2^2 + 2A_1A_2\cos(2\pi \boldsymbol{b}_{pq} \cdot \boldsymbol{s}_0)}. \tag{17}$$

The eigenvalues of $\boldsymbol{\mathcal{R}}_{pq}(u_{pq}, v_{pq})$ are equal to (see Section 3.1)

$$\lambda_1 = c_0 + |c| \tag{18}$$
$$\lambda_2 = c_0 - |c|.$$

The unnormalized eigenvectors of $\boldsymbol{\mathcal{R}}_{pq}(u_{pq}, v_{pq})$ are equal to (see Section 3.1)

$$\mathbf{x}^{(1)} = [|c| \; c]^T, \tag{19}$$
$$\mathbf{x}^{(2)} = [-|c| \; c]^T.$$

Eq. 3, Eq. 4, Eq. 18 and Eq. 19 implies that

$$\boldsymbol{\mathcal{G}}_{pq}(u_{pq}, v_{pq}) = \begin{bmatrix} V_{pp}^{\mathcal{G}} & V_{pq}^{\mathcal{G}} \\ V_{qp}^{\mathcal{G}} & V_{qq}^{\mathcal{G}} \end{bmatrix} \tag{20}$$
$$= \begin{bmatrix} \frac{c_0+|c|}{2} & \frac{c^*(c_0+|c|)}{2|c|} \\ \frac{c(c_0+|c|)}{2|c|} & \frac{c_0+|c|}{2} \end{bmatrix}.$$

with

$$\boldsymbol{g}_{pq}(u_{pq}, v_{pq}) = \begin{bmatrix} \frac{\sqrt{c_0+|c|}}{\sqrt{2}} \\ \frac{c\sqrt{c_0+|c|}}{\sqrt{2}|c|} \end{bmatrix}. \tag{21}$$

4.2 Three element interferometer

When the sky is equal to $I_{\mathcal{R}}(\boldsymbol{s})$ then the unpolarized visibility matrix $\boldsymbol{\mathcal{R}}_{pqr}(\boldsymbol{b})$ (assuming there is a relation between the uv tracks of the interferometer) for antenna p, q and r will be equal to

$$\boldsymbol{\mathcal{R}}_{pqr}(\boldsymbol{b}) = \begin{bmatrix} V_{pp}^{\mathcal{R}} & V_{pq}^{\mathcal{R}} & V_{pr}^{\mathcal{R}} \\ V_{qp}^{\mathcal{R}} & V_{qq}^{\mathcal{R}} & V_{qr}^{\mathcal{R}} \\ V_{rp}^{\mathcal{R}} & V_{rq}^{\mathcal{R}} & V_{rr}^{\mathcal{R}} \end{bmatrix} \tag{22}$$
$$= \begin{bmatrix} A_1 + A_2 & A_1 + A_2 e^{-2\pi i \boldsymbol{b}_{pq}\cdot\boldsymbol{s}_0} & A_1 + A_2 e^{-2\pi i \boldsymbol{b}_{pr}\cdot\boldsymbol{s}_0} \\ A_1 + A_2 e^{-2\pi i \boldsymbol{b}_{qp}\cdot\boldsymbol{s}_0} & A_1 + A_2 & A_1 + A_2 e^{-2\pi i \boldsymbol{b}_{qr}\cdot\boldsymbol{s}_0} \\ A_1 + A_2 e^{-2\pi i \boldsymbol{b}_{rp}\cdot\boldsymbol{s}_0} & A_1 + A_2 e^{-2\pi i \boldsymbol{b}_{rq}\cdot\boldsymbol{s}_0} & A_1 + A_2 \end{bmatrix},$$

where $\boldsymbol{b} = (u, v)$.
Let ϕ_1, ϕ_2 and ϕ_3 be three positive numbers. For an east-west array there exists a linear relation between the uv tracks associated with antenna p, q and r, which can be formally expressed with

$$\boldsymbol{\Phi}_{pqr} = \begin{bmatrix} 0 & \phi_1 & \phi_2 \\ -\phi_1 & 0 & \phi_3 \\ -\phi_2 & -\phi_3 & 0 \end{bmatrix},$$

and

$$r_{pqr}^{j,k}(\boldsymbol{b}) = A_1 + A_2 e^{-2\pi i \phi_{pqr}^{j,k} \boldsymbol{b}\cdot\boldsymbol{s}_0}, \tag{23}$$

implying that

$$\boldsymbol{\mathcal{R}}_{pqr}(\boldsymbol{b}) = \begin{bmatrix} A_1 + A_2 & A_1 + A_2 e^{-2\pi i \phi_1 \boldsymbol{b}\cdot\boldsymbol{s}_0} & A_1 + A_2 e^{-2\pi i \phi_2 \boldsymbol{b}\cdot\boldsymbol{s}_0} \\ A_1 + A_2 e^{2\pi i \phi_1 \boldsymbol{b}\cdot\boldsymbol{s}_0} & A_1 + A_2 & A_1 + A_2 e^{-2\pi i \phi_3 \boldsymbol{b}\cdot\boldsymbol{s}_0} \\ A_1 + A_2 e^{2\pi i \phi_2 \boldsymbol{b}\cdot\boldsymbol{s}_0} & A_1 + A_2 e^{2\pi i \phi_3 \boldsymbol{b}\cdot\boldsymbol{s}_0} & A_1 + A_2 \end{bmatrix}, \tag{24}$$
$$= \begin{bmatrix} c_0 & c_1^* & c_2^* \\ c_1 & c_0 & c_3^* \\ c_2 & c_3 & c_0 \end{bmatrix}. \tag{25}$$

The matrix $\boldsymbol{\mathcal{R}}_{pqr}(\boldsymbol{b})$ is a rank two Hermitian matrix [1] and therefore its characteristic equation simplifies to

$$|\boldsymbol{\mathcal{R}}_{pqr} - \lambda \boldsymbol{I}| = \lambda^3 + \alpha_1\lambda^2 + \alpha_2\lambda, \tag{26}$$

where

$$\alpha_1 = -3c_0, \tag{27}$$

$$\alpha_2 = 3c_0^2 - |c_1|^2 - |c_2|^2 - |c_3|^2, \tag{28}$$

implying that

$$\lambda_1 = \frac{3c_0 + \sqrt{4|c_1|^2 + 4|c_2|^2 + 4|c_3|^2 - 3c_0^2}}{2} \tag{29}$$

$$= \frac{3c_0 + h}{2}, \tag{30}$$

$$\lambda_2 = \frac{3c_0 - h}{2}, \tag{31}$$

$$\lambda_3 = 0, \tag{32}$$

with

$$h = \sqrt{9(A_1^2 + A_2^2) + \sum_{i=1}^{3} 8A_1A_2\cos(2\pi\phi_i \boldsymbol{b}\cdot\boldsymbol{s}_0) - 6A_1A_2}. \tag{33}$$

Eq. 12 implies that the eigenvector associated with λ_1 is equal to

$$\mathbf{x}^{(1)} = \left[\left(\frac{-(c_0+h)}{2}, c_1^*, c_2^*\right) \times \left(c_2, c_3, \frac{-(c_0+h)}{2}\right)\right]^T \tag{34}$$

$$= \left(\underbrace{\begin{vmatrix} c_1^* & c_2^* \\ c_3 & \frac{-(c_0+h)}{2}\end{vmatrix}}_{|\boldsymbol{\mathcal{A}}|}, \underbrace{\begin{vmatrix} c_2^* & \frac{-(c_0+h)}{2} \\ \frac{-(c_0+h)}{2} & c_2\end{vmatrix}}_{|\boldsymbol{\mathcal{B}}|}, \underbrace{\begin{vmatrix} \frac{-(c_0+h)}{2} & c_1^* \\ c_2 & c_3\end{vmatrix}}_{|\boldsymbol{\mathcal{C}}|}\right)^T. \tag{35}$$

The matrix $\boldsymbol{\mathcal{G}}_{pqr}(\boldsymbol{b})$ is therefore equal to

$$\boldsymbol{\mathcal{G}}_{pqr}(u,v) = \frac{\lambda_1}{||\boldsymbol{\mathcal{A}}||^2 + |\boldsymbol{\mathcal{B}}|^2 + ||\boldsymbol{\mathcal{C}}||^2}\begin{bmatrix} |\boldsymbol{\mathcal{A}}||\boldsymbol{\mathcal{A}}|^* & |\boldsymbol{\mathcal{A}}||\boldsymbol{\mathcal{B}}| & |\boldsymbol{\mathcal{A}}||\boldsymbol{\mathcal{C}}|^* \\ |\boldsymbol{\mathcal{B}}||\boldsymbol{\mathcal{A}}|^* & |\boldsymbol{\mathcal{B}}||\boldsymbol{\mathcal{B}}| & |\boldsymbol{\mathcal{B}}||\boldsymbol{\mathcal{C}}|^* \\ |\boldsymbol{\mathcal{C}}||\boldsymbol{\mathcal{A}}|^* & |\boldsymbol{\mathcal{C}}||\boldsymbol{\mathcal{B}}| & |\boldsymbol{\mathcal{C}}||\boldsymbol{\mathcal{C}}|^* \end{bmatrix}, \tag{36}$$

while $\boldsymbol{g}_{pqr}(\boldsymbol{b})$ equals

$$\boldsymbol{g}_{pqr}(\boldsymbol{b}) = \sqrt{\frac{\lambda_1}{||\boldsymbol{\mathcal{A}}||^2 + |\boldsymbol{\mathcal{B}}|^2 + ||\boldsymbol{\mathcal{C}}||^2}}\begin{bmatrix} |\boldsymbol{\mathcal{A}}| \\ |\boldsymbol{\mathcal{B}}| \\ |\boldsymbol{\mathcal{C}}| \end{bmatrix}. \tag{37}$$

References

[1] Grobler T.L., C.D. Nunhokee, O.M. Smirnov, A.J. van Zyl and A.G. de Bruyn, "Calibration artefacts in radio interferometry – I. Ghost sources in Westerbork Synthesis Radio Telescope data", *Montly Notices of the Royal Astronomical Society*, vol. 439, no. 4, 2014, pp. 4035–4047.

[2] Boonstra A.J. and van der Veen A.J., "Gain calibration methods for radio telescope arrays",*IEEE Transactions on Signal Processing*, vol. 51, no. 1, 2003, pp. 25–38.

[3] Thompson A.R., Moran J.M., and Swenson G.W., "Interferometry and synthesis in radio astronomy", John Wiley & Sons, 2008.

[4] Blinn, J. "Consider the lowly 2 x 2 matrix.", *Computer Graphics and Applications, IEEE* vol. 16, no.2, 1996, pp. 82-88.

[5] Kopp, J., "Efficient numerical diagonalization of hermitian 3×3 matrices", *International Journal of Modern Physics C*, vol, 19 no. 3, 2008, pp. 523-548.