

mjk, 30 November, 1984

~~AT/25.1.1/008~~

The choice of linearly polarized feeds for the AT complicates the overall calibration procedure. Essentially, before we can calibrate the array's polarization characteristics we need to establish a large number of array parameters. (For example, the station coordinates; the clock error; ect). To do this we should observe point sources in the I stokes parameter, which requires us to combine correlator products, which can only be done if the polarization properties are known. (The saving grace in this is that there is a marked hierarchy in the errors and their influence, so that a simple iterative scheme is quite possible).

This note - which is, to a large extent, a reworking of the Komesaroff AT/21.3.1/010 & /011 - addresses this problem, in the form of three questions:

- a. If the polarization calibration had been performed, how would we treat the data. (ie. the correlator products)?
- b. What are the consequences of calibration errors? What limits are placed on the achievable dynamic range, for example?
- c. How might we calibrate?

We show that each antenna can be characterized by 3 matrices, each 2x2; a gain matrix, a cross-polarization matrix and a rotation matrix.

The gain matrix describes the gain amplitude and phase characteristics of each polarization channel.

The cross-polarization matrix contains a number of trigonometric terms based on 4 angles, 2 for each polarization. They relate to the ellipticity and orientation of the antenna/feed.

The rotation matrix involves the parallactic angle, and serves to transform the alt-az. coordinate frame of the antenna to the equatorial frame.

The ON-line processing should produce stokes parameters expressed in the equatorial frame, free of antenna defects - the processing can be simply expressed as 3 matrix operations on the correlator products.

We show that a relatively simple iterative scheme can be set up to calibrate the instrument, each pass producing correction terms to be added to the antenna matrices.

In order to calibrate the pointing and polarization characteristics we need a network of calibration sources. Of preference, these should be point, of known position. At least one source of known flux density is needed in order to place the array on the flux density scale. A polarized source of known position angle is required in order to determine the mean feed rotation angle. The instrumental circular polarization requires a source of zero circular polarization. As Komesaroff noted, the magnitude of a source's polarization can be determined from the observations, independantly of the calibration, so that sources of zero linear polarization can be identified. Such sources may simplify the determination of elevation angle dependance in the polarization properties.

There are 4 sections: a section on notation and algebra; this is intended to simplify the discussion, as we provide a fairly compact notation. Then follow 3 sections, dealing with the 3 questions above.

The basic conclusion is that it will be possible to calibrate the AT adequately for polarization measurements, and that the calibration algorithm is reasonably simple.

A. Notation.

We use bold type UPPER CASE (eg. \mathbf{E}) to represent a matrix (usually 2×2), and bold type, lower case (eg. \mathbf{v}) to represent a vector.

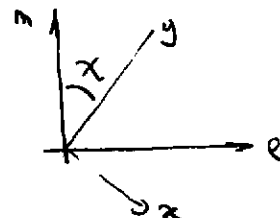
The discussion is essentially based on Weiler (1973), Schwab (1984) and Komesaroff (1984). The intention is to relate the Stokes parameters (as Fourier components) to the correlator products.

a. The electric field:

Let $\mathbf{e} = \{e_1, e_2\}$ be the electric field to be measured, expressed in the equatorial coordinate frame. In the frame of an alt-az. telescope we have:

$$\mathbf{e}_t = \begin{pmatrix} \cos(\chi) & -\sin(\chi) \\ \sin(\chi) & \cos(\chi) \end{pmatrix} \begin{pmatrix} e_1 \\ e_2 \end{pmatrix}$$

$$\mathbf{e}_t = \mathbf{R}_\chi \cdot \mathbf{e}$$



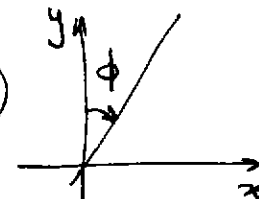
For the compact array, χ is the same for all antennas. (Over 6 km. $\Delta\chi < 10'$, except for the odd source transiting through the zenith).

b. A feed's response to an electric field is characterized by 2 angles. θ and ϕ ; ϕ is a measure of the in-phase cross-polarization, and θ is the quadrature response. (If the cross-polarization is due to the feed alone, ϕ is the position angle of the feed).

The voltage at the feed's terminals can be written:

$$\mathbf{v} = \begin{pmatrix} \cos(\theta) & -i \cdot \sin(\theta) \\ \sin(\theta) & i \cdot \cos(\theta) \end{pmatrix} \begin{pmatrix} \sin(\phi) & \cos(\phi) \\ \cos(\phi) & -\sin(\phi) \end{pmatrix} \begin{pmatrix} e_1 \\ e_2 \end{pmatrix}$$

$$= \mathbf{a} \cdot \mathbf{e}$$



For a horizontally oriented feed, ϕ is essentially 90° , and for a vertical feed, ϕ is close to 0° . It should be noted that in this formulation we are tacitly assuming that there will be no cross-coupling in the receivers, nor in the IF system.

c. We need to allow for the complex gain of the IF channel - call this g . This will encompass the phase introduced by the clock and LO errors, as well as the phase shifts in the Receiver/IF chains.

d. We also need to reckon with a telescope-based phase term arising from the atmosphere/ionosphere. Call it ψ .

The correlator products then take the form:

$$\begin{aligned} S_{ab} &= \langle v_a v_b^* \rangle \\ &= e^{i(\psi_a - \psi_b)} g_a g_b^* (\mathbf{a}_a \cdot \mathbf{e}_a) (\mathbf{a}_b \cdot \mathbf{e}_b)^{*T} \\ &= e^{i(\psi_a - \psi_b)} g_a g_b^* (\mathbf{a}_a) (\mathbf{e}_a \cdot \mathbf{e}_b^*) (\mathbf{a}_b^*{}^T) \end{aligned}$$

The matrix which results from the 2 electric field vectors describes the Stokes parameters; - as fourier components for an interferometer

$$\begin{aligned} \mathbf{e}_a \mathbf{e}_b^* &= \begin{pmatrix} \langle e_x e_x^* \rangle & \langle e_x e_y^* \rangle \\ \langle e_y e_x^* \rangle & \langle e_y e_y^* \rangle \end{pmatrix} \\ &= 1/2 \mathbf{R}\chi_a \begin{pmatrix} I-Q & U-iV \\ U+iV & I+Q \end{pmatrix} \mathbf{R}\chi_b^T \\ &= \mathbf{R}\chi_a \cdot \mathbf{P}_{ab} \cdot \mathbf{R}\chi_b^T \end{aligned}$$

$$\text{where } \mathbf{P}_{ab} = 1/2 \begin{pmatrix} I-Q & U-iV \\ U+iV & I+Q \end{pmatrix}$$

Consider now the 4 correlator products from a single baseline. We use subscripts (a,b) for the antenna, and superscripts (h,v) for the feed orientation. (The "h" feed has a horizontal orientation, and should be parallel to the x-axis).

each will take the form:

$$S_{ab}^{hh} = e^{i(\psi_a - \psi_b)} g_a^h g_b^{h*} \mathbf{a}_a^h (R.P_{ab}.R^T) \mathbf{a}_b^{h*T}$$

The four products can conveniently be combined in a 2x2 matrix:

$$S_{ab} = \begin{pmatrix} S^{hh} & S^{hv} \\ S^{vh} & S^{vv} \end{pmatrix},$$

$$S_{ab} = e^{i(\psi_a - \psi_b)} \mathbf{G}_a \cdot \mathbf{A}_a \cdot (R.P_{ab}.R^T) \cdot \mathbf{A}_b^* \cdot \mathbf{G}_b^{*T} \quad \dots (1)$$

where

$$\mathbf{G}_a = \begin{pmatrix} g_a^h & 0 \\ 0 & g_a^v \end{pmatrix}$$

and

$$\mathbf{A}_a = \begin{pmatrix} \mathbf{a}_a^{hT} \\ \mathbf{a}_a^{vT} \end{pmatrix}$$

B. On-Line reduction of the correlator products.

We intend to convert the correlator products to Stokes parameters automatically, on-line. Thus we need to invert equation (1): $P(\text{est})$ is our estimate of P , derived from the correlator products:

$$P(\text{est}) = R\chi^{-1} \cdot A_s^{-1} \cdot G_s^{-1} \cdot S_{ab} \cdot G_s^{*T-1} \cdot A_s^{*T-1} R\chi \quad \dots (2)$$

Thus the correlator products pass through 3 correction stages: we compensate for gain/phase changes in the IF, then for cross-polarization, and finally for parallactic rotation, in order to obtain the stokes parameters in the equatorial frame.

The 3 correction matrices (R , P , G) have quite different time scales:

R depends on χ , and so must be reevaluated every integration interval;

A should be constant, as it depends on the telescope structure. However, if Parkes is a guide, some elevation dependance may need to be catered for.

G may vary throughout the observation, and will be allowed for:-

A "relative complex gain monitor" will be provided, one for each telescope/frequency band. In essence it will monitor the magnitude of the gain of each IF channel, as well as the relative phase of the two polarization channels. Its output will be a continuous monitor of the 3 quantities:

$$v^v(t) = |v^v|$$

$$v^h(t) = |v^h|$$

$$\rho(t) = \text{phase}(v^h/v^v)$$

where v is the voltage due to a switched noise tube connected to both h and v feeds.

After the calibration exercise we will have established values for g^h and g^v , for some reference values v^h , v^v , and ρ . During the course of the observations we will then

correct g:

$$g^h(t) = g^h(\text{ref}) \cdot (\nu^h(\text{ref})/\nu^h(t)) \cdot e^{i\{\rho(\text{ref}) - \rho(t)\}}$$

$$g^v(t) = g^v(\text{ref}) \cdot (\nu^v(\text{ref})/\nu^v(t))$$

(This assumes that the analog/digital filter question has been resolved, and that ν and ρ truly monitor the signals going to the correlator). The accuracy to which the gain and phase can be established will depend, of course, on the bandwidth and integration time; thus it depends on the time-scale of the receiver variations. Accuracies of order 1% and 1° in 1 second integration at 160 MHz bandwidth requires a calibration noise step of 3 K. Phase variations between telescopes are neither monitored nor corrected by this machinery; calibration source observations, self-calibration analysis and redundancy configurations will be required.

C. Consequences of calibration error.

We will estimate the stokes parameters from the correlator products by inverting equation (2). But since the quantities involved in this operation are imperfectly known (the basic calibration problem), our estimate of P will be subject to error. In this section we attempt to assess the consequences of the calibration errors, and to estimate the limitations to the achievable dynamic range in the final maps. Specifically, we examine the consequences of errors in G , θ , and ϕ . Our estimate of P can be written in the form:

$$P(\text{est}) = e^{i(\psi_a - \psi_b)} C_a \cdot P_{ab} \cdot C_b^T$$

where C (a matrix which ideally should be the identity):

$$C_a = R\chi^{-1} \cdot A_a^{-1} \cdot G_a^{-1} \cdot G_a \cdot A_a \cdot R\chi$$

The calibration based quantities are indicated by (').

We expand C , (on the assumption that the errors in G and A are modest), to provide expressions relating the estimated stokes parameters to the actual values and to the calibration errors. (The gory details are given in the appendix. This formulation assumes that there is, in fact, little cross-polarization. This is probably a valid approximation, as the design intentions are that the cross-polarization should be less than 30 db.)

we put $(g/g') = 1 + \epsilon$; $\zeta = \Delta\phi - i\Delta\theta$. Then:

$$\begin{aligned} 2I(\text{est}) = 2I & & (3) \\ & + I \cdot ((\epsilon^v + \epsilon^h)_a + (\epsilon^v + \epsilon^h)_b^*) \\ & + Q \cdot (((\epsilon^v - \epsilon^h)_a + (\epsilon^v - \epsilon^h)_b^*) \cos(2\chi) \\ & \quad - ((\zeta^v - \zeta^h)_a + (\zeta^v - \zeta^h)_b^*) \sin(2\chi)) \\ & + U \cdot (((\epsilon^v - \epsilon^h)_a + (\epsilon^v - \epsilon^h)_b^*) \sin(2\chi) \\ & \quad + ((\zeta^v - \zeta^h)_a + (\zeta^v - \zeta^h)_b^*) \cos(2\chi)) \\ & - iV \cdot ((\zeta^v + \zeta^h)_a - (\zeta^v + \zeta^h)_b^*) \end{aligned}$$

$$\begin{aligned}
2Q(\text{est}) &= 2Q & (4) \\
&+ I.(((\varepsilon^v - \varepsilon^h)_a + (\varepsilon^v - \varepsilon^h)_b^*) \cos(2\chi) \\
&\quad - ((\zeta^v - \zeta^h)_a + (\zeta^v - \zeta^h)_b^*) \sin(2\chi)) \\
&+ Q. ((\varepsilon^v + \varepsilon^h)_a + (\varepsilon^v + \varepsilon^h)_b^*) \\
&+ U. ((\zeta^v + \zeta^h)_a + (\zeta^v + \zeta^h)_b^*) \\
&- iV.(((\varepsilon^v - \varepsilon^h)_a - (\varepsilon^v - \varepsilon^h)_b^*) \sin(2\chi) \\
&\quad + ((\zeta^v - \zeta^h)_a - (\zeta^v - \zeta^h)_b^*) \cos(2\chi))
\end{aligned}$$

$$\begin{aligned}
2U(\text{est}) &= 2U & (5) \\
&+ I.(((\varepsilon^v - \varepsilon^h)_a + (\varepsilon^v - \varepsilon^h)_b^*) \sin(2\chi) \\
&\quad + ((\zeta^v - \zeta^h)_a + (\zeta^v - \zeta^h)_b^*) \cos(2\chi)) \\
&- Q. ((\zeta^v + \zeta^h)_a + (\zeta^v + \zeta^h)_b^*) \\
&+ U. ((\varepsilon^v + \varepsilon^h)_a + (\varepsilon^v + \varepsilon^h)_b^*) \\
&+ iV.(((\varepsilon^v - \varepsilon^h)_a - (\varepsilon^v - \varepsilon^h)_b^*) \cos(2\chi) \\
&\quad - ((\zeta^v - \zeta^h)_a - (\zeta^v - \zeta^h)_b^*) \sin(2\chi))
\end{aligned}$$

$$\begin{aligned}
2iV(\text{est}) &= 2iV & (6) \\
&+ I. ((\zeta^v + \zeta^h)_a - (\zeta^v + \zeta^h)_b^*) \\
&- Q.(((\varepsilon^v - \varepsilon^h)_a - (\varepsilon^v - \varepsilon^h)_b^*) \sin(2\chi) \\
&\quad + ((\zeta^v - \zeta^h)_a - (\zeta^v - \zeta^h)_b^*) \cos(2\chi)) \\
&+ U.(((\varepsilon^v - \varepsilon^h)_a - (\varepsilon^v - \varepsilon^h)_b^*) \cos(2\chi) \\
&\quad - ((\zeta^v - \zeta^h)_a - (\zeta^v - \zeta^h)_b^*) \sin(2\chi)) \\
&+ iV. ((\varepsilon^v + \varepsilon^h)_a + (\varepsilon^v + \varepsilon^h)_b^*)
\end{aligned}$$

(The term $e^{i(\psi_a - \psi_b)}$ which we have omitted in equations 3 to 6 is common to all of them; that is, it is a common scale factor.)

The complex relative gain monitor should allow us to correct variations in $\{\varepsilon\}$ that occur during the course of an observation, to well below the 1% level. However, calibrating $\{\varepsilon\}$ and $\{\zeta\}$ is a different (and harder) question. The requirements can be estimated.

In general, $Q \lesssim 0.1 I$ and $V \lesssim 0.01 I$. Hence, to measure Q with 10 % accuracy ($0.1 \pm 0.01 I$), we require $\{\epsilon\}$ and $\{\zeta\}$ to be below the 1% level.

Similarly, to measure V with 50% accuracy ($0.01 \pm 0.005 I$), we require $\{\zeta\}$ below the 0.5% level.

In the next section we investigate the prospects of achieving this.

D. How do we calibrate?

The difficulty arises because of the multiplicity of errors which all produce phase errors. We list below the major problems which need to be catered for.

1. Station coordinate errors.
2. Telescope position errors. (Location on the pedestal).
3. Movement of the phase centre as the telescope drives in elevation.
4. Station clock error.
5. Array clock error.
6. Delay tracking error.
7. LO. phase error.
8. Correlator phase error.
9. Differential bandpass phase effects.
10. Gain and phase offsets in the receiver/IF chain
11. Polarization properties of the antenna/feed (θ , ϕ).

The identification of items 1 to 9 is relatively straightforward, provided we have a network of point sources at known position. But we need to combine 2 correlator products in order to obtain I: and this requires $\{\epsilon\}$ to be established.

A possible calibration scenario.

The calibration scheme we propose is an iterative one - although a global, least squares approach could possibly be formulated, it is believed that the scheme outlined here is robust, simple, and lends itself to a gradual upgrading of the accuracy with which the instrumental parameters are known. It has a further advantage that a single data format is sufficient: data recorded for calibration purposes will use exactly the same format as the "real" data: we always record stokes parameters.

The essence of the scheme is contained in equations 3 to 6, which relate the estimated stokes parameters to the actual values and to the errors in the instrumental parameters. Three main stages are required:

1. Preliminary single dish measurements provide a reasonably good first estimate of $\{\epsilon\}$ and $\{\zeta\}$;
2. Observe a number of point sources with the array; use the values of $I(\text{est})$ to establish the array parameters;
3. Track a number of point sources over a range of parallactic angles to provide better polarization calibration.

We can iterate on item 3: the correlator products will be reduced to stokes parameters using the currently available G and A matrices. The calibration process will yield $\{\epsilon\}$ and $\{\zeta\}$ with which we can improve the G and A . This scheme also lends itself well to the proposed calibration source file: we could observe a number of calibration sources over a period of time and, providing that they all used the same G and A , we could average the $\{\epsilon\}$ and $\{\zeta\}$.

C.1. Single dish measurements.

We use equations 3 to 6, in a single-dish mode; ie. determine $S_{..}$.

- a. Observe some unpolarized sources, to obtain $|g^h|$ and $|g^v|$. It should be possible to obtain accuracies in the few % range.
- b. Observe polarized sources to obtain the phase of (g^h/g^v) . The important thing here is that there be polarization - we then look at the cross-correlation products; (S^{vh}/S^{hv}) for example, if we were to use the correlator in a "single antenna" mode. Since we cannot rotate the feeds, we should track the source over a range of χ , to ensure that the source is indeed polarized, and that large errors in $\{\zeta\}$ have not resulted in I entering the calculation.
- c. Polarized sources with known position angle can define the mean ϕ . (Otherwise, we can only determine $\Delta\phi$.)
- d. Return to the unpolarized sources to determine $(\phi^v - \phi^h)$ and $(\theta^v + \theta^h)$

It should be possible to determine 3 angles (phase (g^h/g^v) , $(\theta^v + \theta^h)$, and $(\phi^v - \phi^h)$) to accuracies of order

several degrees. At this stage G and A are adequately specified to allow interferometer products to be combined, yielding a good estimate of I .

C.2 Define the array parameters.

Designate one particular antenna as reference - for the phase of the IF/LO system.

We can combine S^v and S^h to yield a good estimate of:

$$I.e^{i(\psi_a - \psi_b)}$$

The array parameters can then be determined, with reference to a network of point source calibrators of accurately known position.

C.3 Polarization calibration.

If we track a point source over a range of parallactic angles, we will find that all four stokes parameters will show a dependance on χ of the form:

$$F = A + B.\cos(2\chi) + C.\sin(2\chi)$$

In equations (3) and (6) we make use of the terms independent of χ : the "A" term is a function of I and V ; with error terms of comparable magnitude, we will neglect the dependance on V . Similarly, from equations (4) and (5) we will use the "B" and "C" terms, which also depend on I and V alone.

From (3) we determine the Real part of $(\epsilon^v + \epsilon^h)$; we can determine all the Imaginary parts (the phase), but we need to specify a reference antenna. (In effect, only the phase difference enters).

From (6) we obtain the imaginary part of $(\zeta^v + \zeta^h)$; thus, we determine the angle θ , the quadrature cross-coupling. The Real part (the in-phase component) can be determined, once a reference direction is set.

In a similar manner, equations (4) and (5) allow us to measure the difference between the v and h feeds, thus

providing all the factors.

It should be emphasized that the calibration procedure requires us to provide 3 bits of information : the flux density - to set the scale of the data, the position angle of the polarization, and the magnitude of circular polarization, to set the instrumental circular polarization. In addition, we assume that the source is at the tracking centre of the interferometers.

In some more detail:

1. From equation (3):

$$A_{ab} = I + I((\epsilon^v + \epsilon^h)_a + (\epsilon^v + \epsilon^h)_b^*) / 2$$

$$A_{ab} + A_{bc} - A_{ca}^* = ((\epsilon^v + \epsilon^h)_b + (\epsilon^v + \epsilon^h)_c^*) / 2 + I$$

$$= (\text{Real part of } (\epsilon)_b) + I$$

We then set the mean phase of the reference antenna (eg. b) to zero, so that from each A we can solve for the mean phase of the remaining antennas.

2. Equation (6) is treated similarly.

3. From (4) and (5) we get $(\epsilon^v - \epsilon^h)$ and $(\zeta^v - \zeta^h)$.

We could solve for "B" and "C" separately in (4) and (5), or, somewhat more tidily, form:

$$E_1 = \cos(2\chi).(4) + \sin(2\chi).(5) \quad \dots (4')$$

$$E_2 = -\sin(2\chi).(4) + \cos(2\chi).(5) \quad \dots (5')$$

The constant part (indip. of χ) in (4') yields $\{\epsilon\}$, while the constant part in (5') yields $\{\zeta\}$.

The values of $\{\epsilon\}$ and $\{\zeta\}$ which result from these operations can be added to the then current set to improve the calibration factors.

C.4 What accuracy will be achievable?

We can identify a number of determinants of the accuracy of $\{\epsilon\}$ and $\{\zeta\}$:

1. The quality of the calibrators. Are they at the assumed position, of the adopted flux density, linear polarization position angle, and with zero circular polarization?
2. The phase noise introduced by the atmosphere.
3. Instrumental errors.

At the moment there is no obvious answer; presumably it will be a bootstrap operation. 1% and 1° if we are lucky/careful?

References

- Komesaroff, M (1984) AT/23.1.1/010 and 011
Schwab, F (1984) VLB mem. #337
Weiler, K (1973) A&A. 26, p.403

APPENDIX A

The correction matrix (C)

$$C_s = R\chi^{-1} \cdot A_s^{-1} \cdot G_s^{-1} \cdot G_s \cdot A_s \cdot R\chi$$

1.

$$G = \begin{pmatrix} g^h & 0 \\ 0 & g^v \end{pmatrix}$$

$$G^{-1} \cdot G = \begin{pmatrix} 1+\epsilon^h & 0 \\ 0 & 1+\epsilon^v \end{pmatrix}$$

(An enthusiast may wish to allow also for the determinant, which will be slightly > 1).

2.

$$A = \begin{pmatrix} A_{11} & A_{21} \\ A_{12} & A_{22} \end{pmatrix}$$

$$A_{11} = (\cos(\theta)\sin(\phi) - i\sin(\theta)\cos(\phi))^h$$

$$A_{21} = (\cos(\theta)\cos(\phi) + i\sin(\theta)\sin(\phi))^h$$

$$A_{12} = (\cos(\theta)\sin(\phi) - i\sin(\theta)\cos(\phi))^v$$

$$A_{22} = (\cos(\theta)\cos(\phi) + i\sin(\theta)\sin(\phi))^v$$

Now $\theta^h \sim \theta^v \sim 0$; and
 $\phi^h \sim \pi/2$; $\phi^v \sim 0$, so that:

$$\mathbf{A} \sim \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

and,

$$\delta \mathbf{A} \sim \begin{pmatrix} 0 & -\Delta\phi + i\Delta\theta \\ \Delta\phi + i\Delta\theta & 0 \end{pmatrix}$$

put $\xi = \Delta\phi - i\Delta\theta$, so that:

$$(\mathbf{A}'^{-1} \cdot \mathbf{G}'^{-1} \cdot \mathbf{G} \cdot \mathbf{A}) = \begin{pmatrix} 1 + \varepsilon^h & -\xi^h \\ \xi^v & 1 + \varepsilon^v \end{pmatrix}$$

3.

Then, $\mathbf{C} = \mathbf{1} + \mathbf{R}^{-1} \cdot (\delta \mathbf{G} + \delta \cdot \mathbf{A}) \cdot \mathbf{R}$

and

$$\mathbf{P}(\text{est}) = \mathbf{P} + \mathbf{R}^{-1} \cdot (\delta \mathbf{G} + \delta \cdot \mathbf{A}) \cdot \mathbf{R} \cdot \mathbf{P} + \mathbf{P} \cdot \mathbf{R}^T \cdot (\delta \mathbf{G} + \delta \cdot \mathbf{A})^T \cdot \mathbf{R}$$