Astron. Astrophys. 342, 201–211 (1999)

ASTRONOMY
AND
ASTROPHYSICS

Integration of the radiative transfer equation for polarized light: the exponential solution

M. Semel and A. López Ariste

DASOP, URA 326, Observatoire de Paris, Section de Meudon, France

Received 18 March 1998 / Accepted 6 November 1998

Abstract. The radiative transfer equation (RTE) for polarized light accepts a convenient exponential solution when the absorption matrix commutes with its integral. We characterize some of the matrix depth variations which are compatible with the commutation condition. Eventually the vector solution may be diagonalized and one may obtain four independent scalar solutions with four optical depths, complex in general. When the commutation condition is not satisfied, one must resort to a determination of an appropriate evolution operator, which is shown to be well determined mathematically, but whose explicit form is, in general, not easy to apply in a numerical code. However, we propose here an approach to solve a general case not satisfying the commutation condition.

Key words: polarization – radiative transfer – methods: numerical – Sun: magnetic fields – stars: magnetic fields

1. Introduction

The use of spectropolarimetry and Zeeman effect to measure magnetic fields in the sun and stars, requires a transfer theory for polarized light, in the presence of a magnetic field. Such transfer equation was first written in the pioneering paper by Unno (1956), where the effect of anomalous dispersion was not considered yet. Rachkowsky (1962; 1967) was the first to include it, and obtained a general transfer equation for polarized light,

$$\frac{d}{dz}\boldsymbol{I} = -\mathsf{K}\boldsymbol{I} + \boldsymbol{J}, \tag{1}$$

in terms of the Stokes parameters I, Q, U and V presented as the components of vector $\boldsymbol{I}$. This equation *resembles* the transfer equation for the scalar case when polarization is ignored. The scalar intensity is replaced by the Stokes vector $\boldsymbol{I}$. The emision term is now a four component vector $\boldsymbol{J}$, and the scalar absorption coefficient becomes a 4×4 matrix K which describes absorption (including anomalous dispersion) in the presence of Zeeman effect. The variable z parameterizes the light path. The transformation from a single equation to a system of four equations (abridged in a vectorial form) changes drastically the nature of the problem and no general explicit analytical solution has been proposed so far.

The first particular solution was obtained by Unno, who applied the equation to the case of a Milne-Eddington atmosphere. Rachkowsky (1967), after including anomalous dispersion, solved the RTE for the same homogeneous atmosphere.

Several procedures were successful in solving the equation numerically (e.g. Beckers & Schröter 1969, Wittman 1974, Rees, Murphy & Durrant 1989, Landi Degl'Innocenti 1976) even for the general case, by subdividing the atmosphere into numerous layers. The lasts are chosen optically thin in the upper atmosphere and eventually an Unno–Rachkowsky solution is taken for the deepest layer, down to optical depth infinity. Difficulties were encountered as well, depending on the particular method chosen to solve the radiative transfer in each individual sublayer. An universal technique, like the Runge–Kutta method, may not be the best approach for a particular case, mainly because of different scales of variation. The mathematical justifications are often not rigourous and numerical tests are always necessary.

The advantage of an eventual analytical solution is obvious. But, up to now, they have been always restricted to homogeneous atmospheric models where only the LTE source function was depth dependent. The constant K matrix has been handled with different mathematical techniques: for instance Kjeldseth Moe (1968) and Stenflo(1971; 1994) used K-diagonalization, van Ballegooijen(1985) preferred Jones calculus.

Now, for the solar case, variations with depth of both the thermodynamical parameters describing the atmosphere and the magnetic field, are not negligible. For instance, Ruiz Cobo & del Toro Iniesta(1992; 1994), and del Toro Iniesta & Ruiz Cobo (1996; 1996) using numerical inversion of the observed Stokes profiles have confirmed the need for inhomogeneous models (see Collados et al. 1994, Westendorp Plaza et al. 1997a,1997b,1997c, and see also del Toro Iniesta & Ruiz Cobo 1996a for a review). All these works raise the interest in analytical methods dealing with non-homogeneous atmospheric models, i.e. with non-constant K matrices.

At first sight, the RTE for polarized light does not seem more complicated to solve than its scalar equivalent (RTE for pure light intensity, with no polarisation):

$$\frac{d}{dz}I = -\kappa I + J,$$

where as usual, κ is the absorption coefficient, I and J are the scalar intensity and emission function respectively and z the geometrical path. This equation has an explicit formal solution, for the layer $z_0 < t < z$:

$$I(z) = \int_{z_0}^{z} e^{-\int_t^z \kappa dt'} J(t) dt + e^{-\int_{z_0}^z \kappa dt'} I(z_0). \quad (2)$$

A direct extrapolation of this scalar solution would yield:

$$\boldsymbol{I}(z) = \int_{z_0}^{z} e^{-\int_t^z \mathsf{K} dt'} \boldsymbol{J}(t) dt + e^{-\int_{z_0}^z \mathsf{K} dt'} \boldsymbol{I}(z_0). \quad (3)$$

Note that, in this expression, all the mathematical operations involving matrices are well defined. For instance, the exponential of a square matrix is defined as

$$e^{\mathsf{L}} = \mathbb{1} + \mathsf{L} + \frac{1}{2!}\mathsf{L}^2 + \frac{1}{3!}\mathsf{L}^3 + \dots \quad (4)$$

Now, just one difference arises between algebras using matrices and scalars: while two scalars always commute, that is not true in general for two matrices:

$$\mathsf{KL} \neq \mathsf{LK}.$$

This difference becomes important when treating for example the derivative of a power of a matrix: it is not true in general that

$$\frac{d}{dz}\mathsf{L}^n = -n\mathsf{K}\mathsf{L}^{n-1},$$

where we have used (and so we will do hereafter) that

$$\frac{d}{dz}\mathsf{L} = -\mathsf{K}.$$

As a further consequence of the non-commutativity of matrices, we have that

$$\frac{d}{dz}e^{\mathsf{L}} \neq -\mathsf{K}e^{\mathsf{L}},$$

in complete contradiction with the scalar case.

In short, Eq. (3) is a solution of the polarized RTE, Eq. (1), when the commutation condition,

$$[\mathsf{K}, \mathsf{L}] = \mathsf{KL} - \mathsf{LK} = 0, \quad (5)$$

holds. Under this assumption, the previous expressions recover an usual scalar appearance, and therefore we are permitted to write a scalar-like formal solution. It is interesting to note that this condition does not imply a constant absorption matrix. In the following sections we will show how to incorporate variations of K with optical depth. Indeed, Landi Degl'Innocenti & Landi Degl'Innocenti (1981; 1985) have already shown how to handle matrices of the form

$$\mathsf{K} = \mathsf{K}' f(z)$$

with K' being a constant matrix. Although not explicitly said in these papers, it is obvious that here K satisfies condition (5).

For the more general case when the commutation condition (5) does not hold, Landi Degl'Innocenti & Landi Degl'Innocenti (1985) have derived a formal solution for the RTE:

$$\boldsymbol{I}(z) = \int_{z_0}^{z} \mathsf{O}(z, z')\boldsymbol{J}(z')dz' + \mathsf{O}(z, z_0)\boldsymbol{I}(z_0), \quad (6)$$

where $\mathsf{O}(z, z')$ is the *evolution operator*, a new 4×4 matrix which obeys the homogeneous equation

$$\frac{d}{dz}\mathsf{O}(z, z_0) = -\mathsf{K}(z)\mathsf{O}(z, z_0), \quad (7)$$

with initial condition

$$\mathsf{O}(z_0, z_0) = \mathbb{1},$$

where $\mathbb{1}$ represents the 4×4 identity matrix. Note that when solution (3) applies, the evolution operator takes an explicit form, namely:

$$\mathsf{O}(z, z') = \exp(\mathsf{L}(z, z')) = \exp\left(-\int_{z'}^{z} \mathsf{K} dt\right) \quad (8)$$

A general method to solve equations of the form of Eq. (7) for linear operators has already been given by Magnus (1954). In this remarkable paper an exponential expression is proposed:

$$\mathsf{O}(z, z') = \exp\left(\Omega(z - z')\right),$$

were the exponent $\Omega(z - z')$ is given by an infinite series.

Eq. (8) turns out to be Magnus' expression when only the first term in the infinite series is kept (indeed the only non–zero one when condition (5) holds).

We now discuss three existing options to solve Eq. (1):

1. Constant matrix assumption. Condition (5) is inmediately satisfied and analytical solutions were found (Unno, 1956; Rachkowsky, 1967).
2. Multi-layer techniques. The atmosphere is considered as made up by numerous successive layers. A crude assumption on the radiative transfer in each optically thin layer is then advanced, and leads to a procedure of numerical integration expected by intuition to converge to the exact solution. A formal proof of convergence was not given, but the numerical tests were indeed satisfactory. See for instance Rees (1987), Rees et al.(1989), Ruiz Cobo & del Toro Iniesta (1992), del Toro Iniesta & Ruiz Cobo (1996).
3. Magnus' solution. By applying linear algebra one can treat the general case, with non commuting matrices (Magnus, 1954).

The constant matrix technique, method (1), is not possible when one wants to abandon the homogeneous magnetic field and atmosphere assumptions. Next, poor economy is the main drawback of method (2). There is some doubt whether one can determine *a priori* the number of layers necessary for a desired precision. Last but not least, a more analytical insight than what a pure numerical method can give is always desired as well. Moreover, our ultimate purpose is magnetometry of the sun or stars. We want to go beyond the first method, the most used, at present,

but limited to a constant absorption matrix and therefore also constant field. Still we must admit that actual observations will allow us to determine only "little" more than a homogeneous atmosphere model, say, at most the magnetic fields at two or three levels in the atmosphere. It is therefore not "economic" to calculate more than a few layers in the atmosphere.

It is striking that Magnus' solution was published two years before the memorial paper by Unno(1956), the first paper on RTE for polarized light, and as yet it has never been mentioned in the astrophysical literature. It is therefore given in Appendix A. The solution given by Magnus is mathematically exact, but it requires the use of Lie algebra, is not economic and can hardly be used in practical computation. It is mentioned here because it confirms the approach of the present paper and complete it.

Our general strategy is, first, to "satisfy" condition (5) as far as possible by extracting from the absorption matrix everything that commutes with its integral and therefore can easily be integrated according to Eq. (3), as explained in Sect. 2. In Sect. 3, we diagonalise the commutative part of the matrix to allow an efficient integration. Then, in Sect. 4, we treat the residual matrix by an appropriate approximation and thus obtain a semi-analytical solution for an optically *finite* layer with arbitrary depth variations. Eventually we can then borrow the techniques from the multi-layer approach and apply our semi-analytical solution to a few layer model to improve the computation.

A few words on the mathematical space where we are working and where the RTE is to be solved, are in order. Magnetometry concerns the 3D real physical space, where the magnetic field can be represented as a 3D "vector" and all physical parameters of the atmosphere determine the coefficients that enter the radiative transfer equation. The last one is much better calculated in another space. Indeed, we have already entered another 4D geometry: the Minkowski space, where the Stokes' 4–vectors are best described. In this geometry, the norm of a vector $\boldsymbol{I}$ is given by $(I^2-Q^2-U^2-V^2)^{\frac{1}{2}}$. It has particular symmetries and is governed by linear algebra. The elementary operations, like absorption and retardation, are presented by matrices for which commutation relations are of particular importance. When condition (5) holds, an exponential solution, scalar like, to a linear equation can easily be derived. Otherwise, we have to turn to Magnus' exponential solution.

The main difficulties originate from the fact that only few variables are **explicitly** common to the "two spaces". Typically scalar variables like z, κ_c, κ_l (see Sect. 2. for their definitions) will appear in both spaces in similar ways. However, rotations of the Stokes reference system will not. Exception is the azimuth rotation. The angle of rotation of the azimuth of the magnetic field in the "real 3D space" corresponds to a rotation in the Minkowski space, but with a double amount. Naturally, when a constant atmosphere is selected in the real space, the corresponding matrix in the Minkowski space will be constant as well. On the other hand some rotation in the Minkowski space may be much easier to handle. For instance, one may find convenient to use generalized Stokes vectors expressed in terms of elliptic states of polarization. Transformations from one set of Stokes reprensentation to another are expressed simply as rotations in the Minkowski space. Except in some limiting cases it is not possible to translate these angles in terms of angles in the physical space. At the same time, the highly non linear relations between magnetic field and the entries of the absorption matrix cannot in general be simplified. Thus, while the RTE can be solved for a given depth variation of the absorption matrix, we cannot, in general, recover analytically the corresponding variation of the magnetic field. We anticipate that numerical methods can overcome this difficulty and profite from the analytical solution in the Minkowski space to treat the depth variations of the magnetic fields and improve both the economy and the precision of the calculations. These considerations apply as well to all other atmospheric conditions, like temperature, pressure, velocity etc.

In some particular cases, the relations between variables in the Minkowski and real spaces may become simplified. For instance, in absence of absorption of linear polarization, whether in the pure longitudinal magnetic field, or alternatively for particular Zeeman patterns, free of linear polarization. Also for the case when all Zeeman components are separated, simple relations hold as will be discussed in the corresponding sections.

2. Transformation of matrix K

We rewrite the transfer equation as

$$\frac{d}{dz}\boldsymbol{I_0} = -\mathsf{K}_0\boldsymbol{I_0} + \boldsymbol{J_0}$$

And, K_0 being invertible, we can define the source function vector, either LTE or not,

$$\boldsymbol{S_0} = \mathsf{K}_0^{-1}\boldsymbol{J_0},$$

so that the transfer equation reads

$$\frac{d}{dz}\boldsymbol{I_0} = -\mathsf{K}_0\left(\boldsymbol{I_0} - \boldsymbol{S_0}\right). \tag{9}$$

Matrix K_0 can be decomposed as follows

$$\mathsf{K}_0 = \kappa_l(z)\begin{pmatrix} 0 & b\cos 2\phi & b\sin 2\phi & c \\ b\cos 2\phi & 0 & \gamma_\circ & -\beta\sin 2\phi \\ b\sin 2\phi & -\gamma_\circ & 0 & \beta\cos 2\phi \\ c & \beta\sin 2\phi & -\beta\cos 2\phi & 0 \end{pmatrix} + (g\kappa_l(z) + \kappa_c(z))\mathbb{1} \tag{10}$$

Where $\kappa_l(z)$ and $\kappa_c(z)$ are the usual scalar absorption coefficients: the selective (at line center) and the continuum one, respectively; ϕ is the azimuth angle of the magnetic field, relative to a fixed reference system, and $\mathbb{1}$ is the 4×4 identity matrix.

This is the general symmetry of K_0; the meaning of parameters g, b, c, β,and $\gamma_\circ$ can be found by comparing expression (10) with the corresponding ones in Landi Degl'Innocenti & Landi Degl'Innocenti(1981; 1985), Rees (1987) or Kawakami(1983).

We can simplify this matrix by rotating it an angle 2ϕ in the plane Q–U. That is, we introduce a rotation matrix

$$\mathsf{R}_1 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos 2\phi & \sin 2\phi & 0 \\ 0 & -\sin 2\phi & \cos 2\phi & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

and its inverse R_1^{-1}, and we apply them to K_0 to obtain

$$\mathsf{K}_1' = \mathsf{R}_1\mathsf{K}_0\mathsf{R}_1^{-1} = \kappa_l(z)\begin{pmatrix} 0 & b & 0 & c \\ b & 0 & \gamma_0 & 0 \\ 0 & -\gamma_0 & 0 & \beta \\ c & 0 & -\beta & 0 \end{pmatrix} + \\ +(g\kappa_l(z)+\kappa_c(z))\mathbb{1}. \quad (11)$$

Applying this transformation to Eq. (9) we obtain:

$$\mathsf{R}_1\frac{d}{dz}\boldsymbol{I_0} = -\left(\mathsf{R}_1\mathsf{K}_0\mathsf{R}_1^{-1}\right)\mathsf{R}_1\boldsymbol{I_0} + \mathsf{R}_1\boldsymbol{J_0} = -\mathsf{K}_1'\boldsymbol{I_1} + \boldsymbol{J_1} \quad (12)$$

where

$$\boldsymbol{I_1} = \mathsf{R}_1\boldsymbol{I_0},$$

$$\boldsymbol{J_1} = \mathsf{R}_1\boldsymbol{J_0}.$$

The left hand side of the transformed transfer equation (12) is equal to

$$\frac{d}{dz}(\mathsf{R}_1\boldsymbol{I}_\circ) - (\frac{d}{dz}\mathsf{R}_1)\boldsymbol{I}_\circ,$$

where we note that

$$(\frac{d}{dz}\mathsf{R}_1)\boldsymbol{I_0} = (\frac{d}{dz}\mathsf{R}_1)\mathsf{R}_1^{-1}\mathsf{R}_1\boldsymbol{I_0} = \\ = 2\left(\frac{d\phi}{dz}\right)\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}\boldsymbol{I_1}, \quad (13)$$

so that we can write the transformed transfer equation as

$$\frac{d}{dz}\boldsymbol{I_1} = -\mathsf{K}_1\boldsymbol{I_1} + \boldsymbol{J_1} \quad (14)$$

where

$$\mathsf{K}_1 = \kappa_l(z)\mathsf{N}_1 + (g\kappa_l(z)+\kappa_c(z))\mathbb{1},$$

with

$$\mathsf{N}_1 = \begin{pmatrix} 0 & b & 0 & c \\ b & 0 & \gamma & 0 \\ 0 & -\gamma & 0 & \beta \\ c & 0 & -\beta & 0 \end{pmatrix},$$

where we have introduced

$$\gamma = \gamma_0 - 2\frac{d\phi}{dz}\frac{1}{\kappa_l}. \quad (15)$$

The meaning of the new Stokes reference system is as follows: after the ϕ rotation, the new generalized Stokes parameters Q_1 and U_1, projections of vector $\boldsymbol{I}$ on axis $\mathbf{Q_1}$ and $\mathbf{U_1}$ in the new reference system, still correspond to linear polarization, but Zeeman linear absorption affects Q_1 only (absorption along the $\mathbf{Q_1}$ axis). Faraday rotation may still affect U_1, but with zero absorption. The parameters I_1 and V_1 are unchanged (the correspoding axis $\mathbf{I}$ and $\mathbf{V}$ are not affected by the ϕ rotation). In the real space, the meaning of this rotation is that the reference for the usual definition of the Stokes parameters is taken parallel to the magnetic field for Q. These new axes rotate with the field.

A second simplification is obtained by the use of a new rotation, given by

$$\mathsf{R}_2 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\alpha & 0 & \sin\alpha \\ 0 & 0 & 1 & 0 \\ 0 & -\sin\alpha & 0 & \cos\alpha \end{pmatrix} \quad (16)$$

and its inverse ${\mathsf{R}_2}^{-1}$, where

$$\cos\alpha = \frac{b}{\sqrt{b^2+c^2}} \\ \sin\alpha = \frac{c}{\sqrt{b^2+c^2}}. \quad (17)$$

By applying it to the matrix K_1 we obtain

$$\mathsf{K}_2' = \mathsf{R}_2\mathsf{K}_1\mathsf{R}_2^{-1} = \kappa_l(z)\begin{pmatrix} 0 & q & 0 & 0 \\ q & 0 & p' & 0 \\ 0 & -p' & 0 & r' \\ 0 & 0 & -r' & 0 \end{pmatrix} + \\ +(g\kappa_l(z)+\kappa_c(z))\mathbb{1}, \quad (18)$$

where

$$q = \sqrt{b^2+c^2}, \\ p' = \frac{\gamma b - \beta c}{q}, \\ r' = \frac{\gamma c + \beta b}{q}. \quad (19)$$

For the singular case $q = 0$, we adopte $\alpha = 0$, $p' = \gamma$ and $r' = \beta$.

The meaning of the new Stokes reference system is as follows: after the α rotation, the new generalized Stokes parameters Q_2 and V_2, projections of $\boldsymbol{I_1}$ on axis $\mathbf{Q_2}$ and $\mathbf{V_2}$ correspond to elliptic polarizations, Zeeman elliptic absorption affects Q_2 only (absorption along axis $\mathbf{Q_2}$). Faraday rotation still affects V_2 (and U_2) but with zero absorption. Parameters I_2 and U_2 are unchanged (the correspoding axis $\mathbf{I}$ and $\mathbf{U}$ are not affected by the α rotation). Note that α is wavelength dependent and therefore the rotation in the Minkowski space is not constant with λ!

In the real space, the meaning of this rotation is not any longer as simple as before. However, note that the most general state of polarisation is elliptic!! At each λ we can determine the ellipse of polarisation absorbed by the Zeeman effect. We then choose it as axis $\mathbf{Q_2}$. The complete new generalized Stokes system follows from transformation matrix R_2. In deriving α, $q = \sqrt{b^2+c^2}$ stands for the total intensity of the elipse of polarisation and $\sin\alpha = \frac{c}{\sqrt{b^2+c^2}}$ is the rate of circular polarisation. Although easy to calculate, α has no simple meaning in terms of the magnetic field, except for the case of a strong field when all the Zeeman components are completely separated. Then $\alpha = 0$ for the π component, and $\tan\alpha = \pm\frac{2\cos\theta}{\sin^2\theta}$ for the σ components, where θ is the inclination angle of the magnetic field.

We repeat here all the steps made for first rotation R_1, defining the new transformed Stokes and emission vectors

$$\begin{aligned} \boldsymbol{I_2} &= \mathsf{R}_2 \boldsymbol{I_1} \\ \boldsymbol{J_2} &= \mathsf{R}_2 \boldsymbol{J_1}, \end{aligned} \tag{20}$$

and calculating the term

$$\left(\frac{d}{dz}\mathsf{R}_2\right)\mathsf{R}_2^{-1} = \left(\frac{d\alpha}{dz}\right)\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix} \tag{21}$$

so that we can write the transformed transfer equation as

$$\frac{d}{dz}\boldsymbol{I_2} = -\mathsf{K}_2 \boldsymbol{I_2} + \boldsymbol{J_2}, \tag{22}$$

where

$$\mathsf{K}_2 = \kappa_l(z)\mathsf{N}_2 + (g\kappa_l(z) + \kappa_c(z))\mathbb{1},$$

with

$$\mathsf{N}_2 = \begin{pmatrix} 0 & q & 0 & 0 \\ q & 0 & p' & -s \\ 0 & -p' & 0 & r' \\ 0 & s & -r' & 0 \end{pmatrix},$$

where we have introduced a new parameter

$$s = +\frac{d\alpha}{dz}\frac{1}{\kappa_l}.$$

This parameter s, a new non-zero entry in K_2, makes it a little more complicate than before. A new transformation is necessary if we want to obtain a simpler matrix like K_2'. The way for this simplification is a third rotation R_3 given by

$$\mathsf{R}_3 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \cos\xi & -\sin\xi \\ 0 & 0 & \sin\xi & \cos\xi \end{pmatrix}, \tag{23}$$

with

$$\cos\xi = \frac{p'}{\sqrt{p'^2 + s^2}} \qquad \sin\xi = \frac{s}{\sqrt{p'^2 + s^2}}. \tag{24}$$

The same mathematical steps of the two previous rotations are repeated for R_3. We pass directly to the final expression for the transfer equation:

$$\frac{d}{dz}\boldsymbol{I_3} = -\mathsf{K}_3 \boldsymbol{I_3} + \boldsymbol{J_3} \tag{25}$$

where K_3 has the following aspect:

$$\mathsf{K}_3 = \kappa_l(z)\mathsf{N}_3 + (g\kappa_l(z) + \kappa_c(z))\mathbb{1},$$

with

$$\mathsf{N}_3 = \begin{pmatrix} 0 & q & 0 & 0 \\ q & 0 & p & 0 \\ 0 & -p & 0 & r \\ 0 & 0 & -r & 0 \end{pmatrix}$$

where we have defined

$$\begin{aligned} p &= \sqrt{p'^2 + s^2} \\ r &= r' - \frac{d\xi}{dz}\frac{1}{\kappa_l}. \end{aligned} \tag{26}$$

By now the matrix, and consequently the RTE, has been simplified in a general way, without any assumption nor constraint. To proceed to an exponential solution for the transfer equation, we need to ensure that commutation condition (5) holds for K_3. A necessary and sufficient condition for that is a matrix N_3 of the form

$$\mathsf{N}_3 = (\mathsf{N}_3)_0 \cdot f(z) \tag{27}$$

where $(\mathsf{N}_3)_0$ is a constant matrix, and $f(z)$ is any scalar function of z. Let P,Q and R be the integrals of p,q and r. When calculating the commutator of N_3 with its integral $[\mathsf{N}_3, \mathsf{L}_3]$, its only *a priori* non–zero entries are $(Pq - Qp)$ or $(Pr - Rp)$. It is very easy to see that these expressions vanish when 1) $p = 0$, 2) $q = r = 0$ or 3) p,q and r are all proportional to the same scalar function $f(z)$. While the treatement of cases 1) and 2) is straightforward, the general case 3) needs some discussion: we rewrite p,q r as $p_0 f(z)$, $q_0 f(z)$ and $r_0 f(z)$. Matrix $(\mathsf{N}_3)_0$ keeps the appearance of N_3 but with p, q and r substituted by p_0,q_0 and r_0. These new variables to be constant is the **necessary and sufficient condition** for writing an exponential scalar-like solution. A constant K matrix implies 7 constant parameters. Matrix N_3 contains only three parameters q, p and r, but we need to keep constant only the two ratios p/q and r/q to satisfy condition (5). After transformation, only two variables are requested to be constant,i.e., two constraints instead of seven originally: 5 degrees of freedom have been earned. Reviewing the transformation process, those 5 degrees of freedom may be used to treat analytically gradients with depth in azimuth, κ_c and κ_l, angles α and ξ, and $f(z)$. Details about how to do it and its application to a numerical code, are left for a forthcoming paper.

For the sake of demonstration, we discuss the variation of angles α and θ alone in the case of separated Zeeman components. In absence of azimuth variation $p' = 0$ and $p = s$.

For the π Zeeman component, polarisation is purely linear, $\alpha = 0$, and we adopt $\xi = 0$ $p = s = 0$: the integration is straightforward (case 1). For the σ Zeeman components, polarisation is elliptic. Since $p' = 0$, we adopt $\xi = \pi/4$ and $r = r'$.

$$\frac{d\alpha}{dz} = \pm\frac{1}{4}\frac{d(\cos\theta)}{dz}\frac{2}{1 + \cos^2\theta}.$$

Absorption in each σ component is proportional to $q \propto 1 + \cos^2\theta$ and also $r = r' \propto 1 + \cos^2\theta$, θ being the only variable depending on depth. We now suggest $f(z) = 1 + \cos^2\theta$ to match the depth variation of q, r and $p = s = \frac{d\alpha}{dz}\frac{1}{\kappa_l}$ (case 3) with $\xi = \pi/4$ To keep the same depth variation for $\frac{d\alpha}{dz}$, we impose a variation of θ such that

$$\frac{d(\cos\theta)}{dz}\frac{1}{(1 + \cos^2\theta)^2} = \text{constant},$$

and we obtain

$$\frac{\cos\theta}{1+\cos^2\theta} + \arctan(\cos\theta) = 2\,\text{constant}(z - z_0)$$

In general, it will be impossible to interpret α in such a simple way.

3. Diagonalization of the off-diagonal matrix

In the last section, matrix K was simplified as much as possible. All the way to the final form we have seen how to incorporate some atmospheric gradients into the exponential solution, and at the end we have required the commutation condition to still hold with minimum freedom restrictions. We can then calculate the matrix in the exponent of the solution; but we must compute the exponential of this matrix too. The last can be done by using the matrix series (4), but we prefer the comfort of calculating the exponential of scalars. We can achieve this purpose by diagonalizing K_3. In fact, we only need to diagonalize $(\mathsf{N}_3)_0$. We solve the eigenvalue equations for this matrix:

$$(\mathsf{N}_3)_0 \boldsymbol{U}_i = \lambda_i \boldsymbol{U}_i$$

The four eigenvalues λ_i are given by:

$$\lambda_i^4 + \lambda_i^2(r_0^2 + p_0^2 - q_0^2) - q_0^2 r_0^2 = 0 \tag{28}$$

which is a biquadratic equation, so first we solve

$$\lambda_i^2 = \frac{(q_0^2 - r_0^2 - p_0^2) \pm \Delta}{2} \tag{29}$$

where

$$\Delta^2 = (r_0^2 + p_0^2 - q_0^2)^2 + 4q_0^2 r_0^2, \tag{30}$$

and the four eigenvalues are therefore:

$$\lambda_1 = -\sqrt{(q_0^2 - r_0^2 - p_0^2 + \Delta)/2}, \tag{31}$$
$$\lambda_2 = -\lambda_1, \tag{32}$$
$$\lambda_3 = -\sqrt{(q_0^2 - r_0^2 - p_0^2 - \Delta)/2}, \tag{33}$$
$$\lambda_4 = -\lambda_3. \tag{34}$$

We introduce the notations:

$$\theta_i = -\lambda_i \tag{35}$$
$$\delta_i = \lambda_i^2 - q_0^2, \tag{36}$$

noting that

$$\theta_1 = -\theta_2$$
$$\theta_3 = -\theta_4,$$

and that

$$\delta_1 = \delta_2$$
$$\delta_3 = \delta_4.$$

Using this notation, one finds that the corresponding eigenvectors $\boldsymbol{U}_i$ are given by

$$\boldsymbol{U}_i = \frac{1}{\sqrt{2}q_0}\begin{pmatrix} q_0 \\ -\theta_i \\ \delta_i/p_0 \\ r_0\delta_i/(p_0\theta_i) \end{pmatrix}. \tag{37}$$

We may now calculate matrices T and T^{-1} which diagonalize $(\mathsf{N}_3)_0$:

$$\mathsf{T}^{-1} = \frac{1}{q_0 p_0 \sqrt{2}} \times \begin{pmatrix} q_0p_0 & q_0p_0 & q_0p_0 & q_0p_0 \\ -p_0\theta_1 & p_0\theta_1 & -p_0\theta_3 & p_0\theta_3 \\ \delta_1 & \delta_1 & \delta_3 & \delta_3 \\ r_0\delta_1/\theta_1 & -r_0\delta_1/\theta_1 & r_0\delta_3/\theta_3 & -r_0\delta_3/\theta_3 \end{pmatrix} \tag{38}$$

and

$$\mathsf{T} = \frac{1}{\Delta\sqrt{2}} \times \begin{pmatrix} -\delta_3 & \theta_1\delta_3/q_0 & q_0p_0 & \theta_1\theta_3^2 p_0/(r_0q_0) \\ -\delta_3 & -\theta_1\delta_3/q_0 & q_0p_0 & -\theta_1\theta_3^2 p_0/(r_0q_0) \\ \delta_1 & -\theta_3\delta_1/q_0 & -q_0p_0 & -\theta_3\theta_1^2 p_0/(r_0q_0) \\ \delta_1 & \theta_3\delta_1/q_0 & -q_0p_0 & \theta_3\theta_1^2 p_0/(r_0q_0) \end{pmatrix}. \tag{39}$$

These transformation matrices T and T^{-1} just derived can be applied in general, save for those exceptions where they become singular. These particular cases are:

1. The case when $p_0 = 0$.
   To solve this singularity we suggest the following transformation matrix and its inverse as a substitution of the previous ones:

   $$\mathsf{T} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & -1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & i \\ 0 & 0 & i & 1 \end{pmatrix}, \tag{40}$$

   with the resultant diagonalization of $(\mathsf{N}_3)_0$:

   $$\mathsf{T}(\mathsf{N}_3)_0\mathsf{T}^{-1} = \begin{pmatrix} -q_0 & 0 & 0 & 0 \\ 0 & +q_0 & 0 & 0 \\ 0 & 0 & -ir_0 & 0 \\ 0 & 0 & 0 & +ir_0 \end{pmatrix}. \tag{41}$$

2. The case when $q_0 = 0$.
   With no loss of generality one can take $\alpha = 0$ and then R_2 (and subsequently R_3) becomes the identity matrix. For the diagonalization transformation, the eigenvalues are the following:

   $$\lambda_1 = 0, \tag{42}$$
   $$\lambda_2 = 0, \tag{43}$$
   $$\lambda_3 = -i\Delta^{\frac{1}{2}}, \tag{44}$$
   $$\lambda_4 = i\Delta^{\frac{1}{2}}. \tag{45}$$

   Here Δ is reduced to

   $$\Delta = p_0^2 + r_0^2, \tag{46}$$

and the diagonalizing matrix is:

$$\mathsf{T} = \frac{1}{2\Delta}\begin{pmatrix} 2\Delta & 0 & 0 & 0 \\ 0 & 2r_0 & 0 & 2p_0 \\ 0 & p_0 & i\Delta^{\frac{1}{2}} & r_0 \\ 0 & p_0 & -i\Delta^{\frac{1}{2}} & -r_0 \end{pmatrix}. \quad (47)$$

3. The case when $r_0 = 0$, and $q_0^2 \neq p_0^2$.
In this case, $\Delta = q_0^2 - p_0^2$; $\Delta^{\frac{1}{2}}$ is imaginary if $q_0^2 < p_0^2$. The eigenvalues are given as follows:

$$\lambda_1 = i\Delta^{\frac{1}{2}}, \quad (48)$$
$$\lambda_2 = -i\Delta^{\frac{1}{2}}, \quad (49)$$
$$\lambda_3 = 0, \quad (50)$$
$$\lambda_4 = 0, \quad (51)$$

and the transformation matrix

$$\mathsf{T} = \frac{1}{2\Delta}\begin{pmatrix} q_0 & i\Delta^{\frac{1}{2}} & p_0 & 0 \\ q_0 & -i\Delta^{\frac{1}{2}} & p_0 & 0 \\ -2p_0 & 0 & -2q_0 & 0 \\ 0 & 0 & 0 & 2\Delta \end{pmatrix}. \quad (52)$$

4. The case when $r_0 = 0$, and $p_0 = q_0$.
This is the only case where matrix $(\mathsf{N_3})_0$ cannot be diagonalized. It will be reduced by using

$$\mathsf{T} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (53)$$

and its inverse, so that we get

$$\mathsf{T}(\mathsf{N_3})_0\mathsf{T}^{-1} = \begin{pmatrix} 0 & q_0 & 0 & 0 \\ 0 & 0 & q_0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}. \quad (54)$$

Note that the resulting matrix, although not diagonal is simple enough as to have its exponential calculated as

$$\mathrm{e}^{-\int \mathsf{T}\mathsf{N_3}\mathsf{T}^{-1}dt} = \begin{pmatrix} 1 & -Q & Q^2/2 & 0 \\ 0 & 1 & -Q & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad (55)$$

where we have used the notation $Q = \int q dt$.

Back to the general case, we diagonalize (or apply the equivalent transformation in case 4) to the RTE. It is a new transformation of the equation, as were R_1, R_2 and R_3, and as in these 3 rotations, a term of the form

$$\left(\frac{d}{dz}\mathsf{T}\right)\mathsf{T}^{-1}\mathsf{T}\boldsymbol{I_3}$$

appears in the transformed equation. However, as all the entries of $(\mathsf{N_3})_0$, and subsequently of T, are constant, this term is immediately zero, in perfect agreement with the statement that expression (27) is a necessary and sufficient condition to write a scalar–like exponential solution.[1]

For the general case, we define the new generalized Stokes and emission vectors after the diagonalizing transformation for the RTE as

$$\boldsymbol{I}_T = \mathsf{T}\boldsymbol{I_3}$$
$$\boldsymbol{J}_T = \mathsf{T}\boldsymbol{J_3} \quad (56)$$

and rewrite the diagonalized transfer equation as

$$\frac{d}{dz}\boldsymbol{I}_T = -\Lambda\boldsymbol{I}_T + \boldsymbol{J}_T \quad (57)$$

where

$$\Lambda = \kappa_l(z)\mathsf{T}(\mathsf{N_3})_0\mathsf{T}^{-1}f(z) + (g\kappa_l(z) + \kappa_c(z))\mathbb{1}$$

is the diagonalized matrix. For the last singular case, the equivalent reduced matrix should be used.

4. The formal solution in a particular atmosphere

In the previous two sections, we have seen how to transform matrix K in order to be able to write an exponential solution in a convenient form, and to be able to calculate that exponential as easily as the scalar case. We have shown that this procedure applies to absorption matrices whose depth dependence has a few degrees of freedom but still satisfying condition (5). Summing up the last two sections, we can manage gradients in the azimuth ϕ of the magnetic field and in the angles α and ξ, variations in κ_l and κ_c, treat analytically a general function $f(z)$ in the final matrix $\mathsf{N_3}$, and integrate for any emission vector $\boldsymbol{J}$.

In this context we assume an ideal atmosphere satisfying condition (27). We write a formal solution (3) for the diagonalized transfer equation (57):

$$\boldsymbol{I}_T = \int_{z_0}^{z} \mathrm{e}^{-\int_t^z \Lambda(t')dt'}\boldsymbol{J}_T(t)dt + \mathrm{e}^{-\int_{z_0}^z \Lambda(t')dt'}\boldsymbol{I}_T(z_0) \quad (58)$$

Except for the last singular case, we can apply a diagonal matrix Λ, and write its elements as

$$\Lambda_{ii} = \kappa_l\lambda_i f(z) + (g\kappa_l + \kappa_c) \quad (59)$$

where the λ_i's coincide with the eigenvalues of the off-diagonal matrix $(\mathsf{N_3})_0$. We can now write a scalar disentangled solution for each one of the four components of the generalized Stokes vector as

$$(\boldsymbol{I}_T)_i(z) = \int_{z_0}^{z} \mathrm{e}^{-\int_t^z(\kappa_l\lambda_i f(t') + g\kappa_l + \kappa_c)dt'}(\boldsymbol{J}_T)_i(t)dt + \\ + \mathrm{e}^{-\int_{z_0}^z(\kappa_l\lambda_i f(t') + g\kappa_l + \kappa_c)dt'}(\boldsymbol{I}_T)_i(z_0) \quad (60)$$

[1] If we relax condition (27), $\mathsf{N_3}$ can still be diagonalized by substituting p, q and r for p_0, q_0 and r_0 in T and T^{-1}. But this time $\frac{d}{dz}\mathsf{T}$ is not zero. We cannot get rid of this term by mathematical manipulations, the RTE cannot be diagonalized and a scalar-like solution as Eq. (3) cannot be applied.

where we can define four generalized optical depths

$$\tau_i(z, z_0) = \int_{z_0}^{z} (\kappa_l \lambda_i f(t') + g\kappa_l + \kappa_c) dt',$$

complex in general, with which the solutions are

$$(\boldsymbol{I}_T)_i(z) = \int_{z_0}^{z} \mathrm{e}^{\tau_i(z,t)} (\boldsymbol{J}_T)_i(t) dt + \mathrm{e}^{\tau_i(z,z_0)} (\boldsymbol{I}_T)_i(z_0). \quad (61)$$

We can relate now the generalized Stokes vector $\boldsymbol{I}_T$ with the physical Stokes vector $\boldsymbol{I_0}$ by writting in order all the transformations that have been made, i.e.

$$\boldsymbol{I}_T = \mathsf{T}\mathsf{R}_3\mathsf{R}_2\mathsf{R}_1\boldsymbol{I_0} = \mathsf{T}_\mathsf{T}\boldsymbol{I_0} \quad (62)$$

where T_T is the complete transformation, product of all others in the proper order. Evidently T_T is invertible, so that we can make back way from the generalized Stokes vector to the physical one:

$$\boldsymbol{I_0} = \mathsf{T}_\mathsf{T}^{-1}\boldsymbol{I}_T.$$

5. An approximation for a general atmosphere

Model atmospheres are usually not so perfect as to be considered in the cases treated in the last section. In the absence of a general analytical solution for the evolution operator, we must manage those general models numerically. Our strategy is to integrate everything we can and to linearize the rest. For this purpose we borrow a technique from other successful numerical integrators as the well-known DELO (Rees et al., 1989). In what follows we develop this idea.

We want to integrate the transfer equation

$$\frac{d}{dz}\boldsymbol{I} = -\mathsf{K}\boldsymbol{I} + \boldsymbol{J} \quad (63)$$

or alternatively

$$\frac{d}{dz}\boldsymbol{I} = -\mathsf{K}\left(\boldsymbol{I} - \boldsymbol{S}\right). \quad (64)$$

The integration is to be made in the interval $z_1 < z < z_2$ and the atmosphere in the two extreme points z_1 and z_2 is specified, that is, we know $\mathsf{K}(z_1)$ and $\mathsf{K}(z_2)$ and also $\boldsymbol{S}(z_1)$ and $\boldsymbol{S}(z_2)$. The incoming light $\boldsymbol{I}(z_1)$ is also given. We want to obtain the polarized light at z_2: $\boldsymbol{I}(z_2)$. After the last section we already know how to integrate the transfer equation if the atmosphere is characterized by the prescriptions given there. Note that, given the atmosphere at the two points z_1 and z_2, one can always calculate the angles α and ξ and the matrix N_3 at those two levels, and look for a suitable integrable atmosphere in agreement with expression (27) which satisfy the data at z_1 and z_2 as far as possible. So let us approximate our atmosphere between these two levels by this integrable atmosphere, represented by a matrix $\overline{\mathsf{K}}$ and an emission vector $\overline{\boldsymbol{S}}$. We can obtain a solution $\overline{\boldsymbol{I}}(z_2)$ for the equation

$$\frac{d}{dz}\overline{\boldsymbol{I}} = -\overline{\mathsf{K}}(\overline{\boldsymbol{I}} - \overline{\boldsymbol{S}}) \quad (65)$$

taking as initial condition $\overline{\boldsymbol{I}}(z_1) = \boldsymbol{I}(z_1)$. Substraction of Eq. (65) from Eq. (64) results in

$$\frac{d}{dz}(\boldsymbol{I} - \overline{\boldsymbol{I}}) = \mathsf{K}(\boldsymbol{I} - \boldsymbol{S}) - \overline{\mathsf{K}}(\overline{\boldsymbol{I}} - \overline{\boldsymbol{S}}), \quad (66)$$

and upon formal integration:

$$\boldsymbol{I}_2 - \overline{\boldsymbol{I}}_2 = \int_{z_1}^{z_2} \left(\mathsf{K}(\boldsymbol{I} - \boldsymbol{S}) - \overline{\mathsf{K}}(\overline{\boldsymbol{I}} - \overline{\boldsymbol{S}})\right) dt, \quad (67)$$

where $\boldsymbol{I}_2 = \boldsymbol{I}(z_2)$ and $\overline{\boldsymbol{I}}_2 = \overline{\boldsymbol{I}}(z_2)$ given as solution to Eq. (65). This equation reflects the error made under the previous approximation. Following our strategy, once we have solved for the integrable part we linearize the rest. So that we now assume that the right hand side of equation (66) is small and can be linearized in the interval $z_1 < z < z_2$. We define

$$y = \mathsf{K}(\boldsymbol{I} - \boldsymbol{S}) - \overline{\mathsf{K}}(\overline{\boldsymbol{I}} - \overline{\boldsymbol{S}}) \quad (68)$$

and upon linearization we write

$$y = a + b(z - z_1), \quad (69)$$

where

$$a = \mathsf{K}_1(\boldsymbol{I}_1 - \boldsymbol{S}_1) - \overline{\mathsf{K}}_1(\overline{\boldsymbol{I}}_1 - \overline{\boldsymbol{S}}_1). \quad (70)$$

At $z = z_2$, we have $y = a + b(z_2 - z_1)$, and we obtain

$$a + b(z_2 - z_1) = \mathsf{K}_2(\boldsymbol{I}_2 - \boldsymbol{S}_2) - \overline{\mathsf{K}}_2(\overline{\boldsymbol{I}}_2 - \overline{\boldsymbol{S}}_2). \quad (71)$$

To solve Eq. (67) we write

$$\boldsymbol{I}_2 - \overline{\boldsymbol{I}}_2 = \int_{z_1}^{z_2} y dt. \quad (72)$$

And by means of the linearization the last integral becomes

$$a(z_2 - z_1) + \frac{b}{2}(z_2 - z_1)^2$$

so that, substituting a and b by its complete expressions

$$\begin{aligned}\boldsymbol{I}_2 - \overline{\boldsymbol{I}}_2 \;=\; & \frac{z_2 - z_1}{2}\left(\mathsf{K}_2(\boldsymbol{I}_2 - \boldsymbol{S}_2) - \overline{\mathsf{K}}_2(\overline{\boldsymbol{I}}_2 - \overline{\boldsymbol{S}}_2)\right. \\ & \left. +(\mathsf{K}_1 - \overline{\mathsf{K}}_1)\boldsymbol{I}_1 - \mathsf{K}_1\boldsymbol{S}_1 + \overline{\mathsf{K}}_1\overline{\boldsymbol{S}}_1\right). \end{aligned} \quad (73)$$

In this expression everything is already known except for $\boldsymbol{I}_2$ that is precisely what we want to calculate.

A convenient choice of the overlined parameters may render equation (73) simpler. For illustration, let us choose

$$\overline{\mathsf{K}}_1 = \overline{\mathsf{K}}_2 = \mathsf{K}_2$$

and $\overline{\boldsymbol{S}}_1 = \boldsymbol{S}_1$, $\overline{\boldsymbol{S}}_2 = \boldsymbol{S}_2$. We then obtain

$$\begin{aligned}\boldsymbol{I}_2 \;=\; & \overline{\boldsymbol{I}}_2 - \left[1 - \frac{z_2 - z_1}{2}\mathsf{K}_2\right]^{-1} \frac{z_2 - z_1}{2} \\ & \times (\mathsf{K}_2 - \mathsf{K}_1)\,(\boldsymbol{I}_1 - \boldsymbol{S}_1), \end{aligned} \quad (74)$$

a solution for $\boldsymbol{I}(z_2)$. This solution is not exact, its precision depends on how good the linear approximation is. In the limit, we can made the integration interval (z_1, z_2) as small as we want but at the cost of increasing the number of layers. A compromise will be necessary between speed and required precision.

6. Conclusions

The purpose of this paper was first to deepen our understanding of the integration of the RTE for polarised light and next to improve the basis for numerical codes. The main conclusions is: *The fundamental key to solve the RTE of polarised light is the commutation of the absorption matrix and its integral.*

When this commutation condition is satisfied:

1. A scalar-like solution can be proposed to the vector equation.
2. A constant absorption matrix satisfies the commutation requirement, however, it is only a sufficient condition, not necessary. After some elaboration one can show that only two constraints at all are necessary instead of the seven inherent in the fully constant absorption matrix.
3. In general, it will possible to diagonalize the absorption matrix and consequently also the RTE with its vector solution. This results in four scalar equations with four scalar solutions. The variables are no more the usual Stokes parameters, but generalized ones, corresponding to general states of polarisation.
4. The solution being analytical, it is valid for quite thick optical layers. The real numerical application is beyond the scope of this paper.

When the commutation condition does not hold, one can turn to Magnus' solution, described shortly in the appendix. Direct application of Magnus' solution to a numerical code seems immature at present. For a general atmosphere, the numerical strategy proposed is to integrate analytically what we can and approximate the rest, that is:

1. Divide the atmosphere into a reasonably number of layers, so that in each of them the commutation condition is only slightly violated.
2. Approximate the general absorption matrix in each layer by an average that satisfies the commutation condition.
3. Apply the solution developed in this paper using last matrix.
4. Applying an approximation for the residual matrix, eventually the one used in DELO (Rees et al., 1989).

As an objective, we intend to improve the efficiency of integration and inversion codes. This will be a must in treating the abundant data expected from multi–line spectropolarimetric observations to be provided by the French—Italian telescope THEMIS.

Acknowledgements. We gratefully thank J.C. del Toro Iniesta, A. Skumanich and G. Chambe for carefully reading this paper and for their valuable comments.

Appendix A: the general solution to the evolution operator

We intend in this appendix to give a self-contained proof of Magnus' exponential solution to the equation for the evolution operator. The original proof is to be found in Magnus' paper(1954). Here we will give a short but complete proof introducing in the meanwhile the necessary tools used while working with Magnus' expressions and to get acquainted with algebraic manipulations inherent to problems involving non-commutative operators. Magnus makes use of an original technique to manage the derivative of the exponential of a linear operator (a square matrix in our case). He transforms this derivative into an algebraical expression, and uses it to solve the differential equation.

In what follows w, x, y and z stand for such operators also called *Lie elements*. The *Lie product* of two Lie elements x and y is defined by

$$w = [x, y]\,,$$

where w is a new Lie element. This product is usually called *commutator*. Now an abreviated notation is used for the l-fold Lie product by x of y as

$$\{y, x^l\} = [[\ldots[y, \overbrace{x]\ldots, x]}^{l \text{ times}}$$

with $\{y, x^0\} = y$. With this notation, it is easy to show by a straightforward calculation that

$$\mathrm{e}^{-x} w \mathrm{e}^{x} = \sum_{l=0}^{\infty} \frac{1}{l!} \{w, x^l\}, \tag{A1}$$

where we remind that the exponential of an operator x is defined as

$$\mathrm{e}^{x} = \sum_{l=0}^{\infty} \frac{1}{l!} x^l.$$

Following Magnus, we extend the previous notation to a polynom $P(x)$ in an evident form:

$$\{y, P(x)\} = \sum_{l=0}^{\infty} p_l \{y, x^l\} \tag{A2}$$

where

$$P(x) = \sum_{l=0}^{\infty} p_l x^l.$$

And we are ready to demonstrate the first two formulas which we will use hereafter.

Formula 1

$$\mathrm{e}^{-x} \left(y \frac{\partial}{\partial x} \right) \mathrm{e}^{x} = \left\{ y, \frac{\mathrm{e}^{x} - 1}{x} \right\}$$

This formula arises from the straightforward calculation of its left-hand-side. We begin calculating the effect of $y\frac{\partial}{\partial x}$ on the left of the exponential:

$$\begin{aligned} \left(y \frac{\partial}{\partial x} \right) \mathrm{e}^{x} \;=\; & y + \frac{1}{2} (yx + xy) \\ & + \frac{1}{3!} \left(yx^2 + xyx + x^2 y \right) + \ldots \end{aligned}$$

Next we multiply on the left by e^{-x}:

$$\mathrm{e}^{-x}\left(y\frac{\partial}{\partial x}\right)\mathrm{e}^{x} =$$
$$= y + \frac{1}{2}\left(yx - xy\right) + \frac{1}{3!}\left(x^2y - 2xyx + yx^2\right) + \ldots =$$
$$= y + \frac{1}{2}\left[y,x\right] + \frac{1}{3!}\left[\left[y,x\right],x\right] + \ldots = \sum_{l=0}^{\infty}\frac{1}{(l+1)!}\{y,x^l\}$$

Now,

$$\sum_{l=0}^{\infty}\frac{1}{(l+1)!}x^l = \frac{\mathrm{e}^x - 1}{x},$$

so Formula 1 is demonstrated.

Formula 2

$$\left(\left(y\frac{\partial}{\partial x}\right)\mathrm{e}^{x}\right)\mathrm{e}^{-x} = \left\{y, \frac{1-\mathrm{e}^{-x}}{x}\right\}$$

The left-hand-side of Formula 2 is obtained by multiplying the left-hand-side of Formula 1 on the left by e^x and on the right by e^{-x}. If we do the same multiplications in the right-hand-side of Formula 1 we obtain

$$\left(\left(y\frac{\partial}{\partial x}\right)\mathrm{e}^{x}\right)\mathrm{e}^{-x} = \mathrm{e}^{x}\left\{y, \frac{\mathrm{e}^x-1}{x}\right\}\mathrm{e}^{-x},$$

or expanding the right-hand-side

$$\left(\left(y\frac{\partial}{\partial x}\right)\mathrm{e}^{x}\right)\mathrm{e}^{-x} = \sum_{l=0}^{\infty}\frac{1}{(l+1)!}\mathrm{e}^{x}\left\{y,x^l\right\}\mathrm{e}^{-x} =$$
$$\sum_{l=0}^{\infty}\sum_{n=l}^{\infty}\frac{(-1)^{n-l}}{(n-l)!(l+1)!}\left\{y,x^n\right\} = \sum_{n=0}^{\infty}c_n\left\{y,x^n\right\}.$$

where

$$c_n = \sum_{l=0}^{n}\frac{(-1)^{n-l}}{(n-l)!(l+1)!},$$

which can be seen to correspond to the expansion of

$$\left\{y, \frac{1-\mathrm{e}^{-x}}{x}\right\}.$$

Next, Magnus demonstrates what can be called

The Magnus' Inversion Lemma: *Let $P(x)$ and $Q(x)$ be two power series in x which satisfy*

$$P(x)Q(x) = 1.$$

Then each of the equations

$$\{y,P(x)\} = u,\ y = \{u,Q(x)\}$$

is a consequence of the other one. Let $P(x) = \sum_l p_l x^l$ and $Q(x) = \sum_m q_m x^m$, then by hypothesis

$$1 = \sum_{l,m} p_l q_m x^{l+m}.$$

Now we can obtain the following equivalent expressions

$$y = \{y,1\} = \left\{y, \sum_{l,m} p_l q_m x^{l+m}\right\} = \sum_{l,m} p_l q_m \left\{y, x^{l+m}\right\},$$

where we have used notation (A2). Next we can separate indexes and write

$$y = \sum_m q_m \left\{\sum_l p_l \left\{y,x^l\right\}, x^m\right\}$$
$$= \sum_m q_m \left\{\left\{y,P(x)\right\},x^m\right\}.$$

If we suppose now that $\{y,P(x)\} = u$, immediately we obtain that

$$y = \sum_m q_m \left\{u,x^m\right\} = \left\{u,Q(x)\right\}.$$

The inverse implication is completely equivalent.

With the Inversion Lemma and Formula 2, we have all the instruments to solve the equation for the evolution operator

Exponential solution Theorem (Magnus): *Let $\mathbf{K}(t)$ be a known function of t in an associative ring (for our purposes it is a matrix), and let $\mathbf{O}(t)$ be an unknown function (in our case the evolution operator) satisfying*

$$\frac{d\mathbf{O}}{dt} = \mathbf{K}\mathbf{O},\ \mathbf{O}(0) = \mathbf{1}, \tag{A3}$$

where $\mathbf{1}$ is the identity matrix. Then, if certain unspecified conditions of convergence are satisfied, $\mathbf{O}(t)$ can be written in the form

$$\mathbf{O}(t) = \exp\Omega(t)$$

where

$$\frac{d\Omega}{dt} = \left\{\mathbf{K}, \frac{\Omega}{1-\mathrm{e}^{-\Omega}}\right\} = \sum_{n=0}^{\infty}\beta_n\left\{\mathbf{K},\Omega^n\right\} =$$
$$= \mathbf{K} + \frac{1}{2}[\mathbf{K},\Omega] + \frac{1}{12}\left\{\mathbf{K},\Omega^2\right\} \mp \ldots$$

The β_n vanish for $n = 3,5,7,\ldots$, and $\beta_{2m} = (-1)^{m-1}B_{2m}/(2m)!$, where the B_{2m} (for $m = 1,2,3,\ldots$) are the Bernoulli numbers.

Integration of this equation by iteration leads to an infinite series for Ω the first terms of which (up to terms involving 4 $\mathbf{K}$'s) are

$$\Omega(t) = \int_0^t \mathbf{K}(\tau)d\tau + \tfrac{1}{2}\int_0^t\left[\mathbf{K}(\tau),\int_0^\tau\mathbf{K}(\sigma)d\sigma\right]d\tau +$$
$$+\tfrac{1}{4}\int_0^t\left[\mathbf{K}(\tau),\int_0^\tau\left[\mathbf{K}(\sigma),\int_0^\sigma\mathbf{K}(\rho)d\rho\right]d\sigma\right]d\tau +$$
$$+\tfrac{1}{12}\int_0^t\left[\left[\mathbf{K}(\tau),\int_0^\tau\mathbf{K}(\sigma)d\sigma\right],\int_0^\tau\mathbf{K}(\sigma)d\sigma\right]d\tau +$$
$$+\tfrac{1}{8}\int_0^t\left[\mathbf{K}(\tau),\int_0^\tau\left[\mathbf{K}(\sigma),\int_0^\sigma\left[\mathbf{K}(\rho),\int_0^\rho\mathbf{K}(\nu)d\nu\right],d\rho\right]d\sigma\right]\times$$
$$\times d\tau + \tfrac{1}{24}\int_0^t\left[\mathbf{K}(\tau),\int_0^\tau\left\{\mathbf{K}(\sigma),\left(\int_0^\sigma\mathbf{K}(\rho)d\rho\right)^2\right\}d\sigma\right]\times$$
$$\times d\tau + \ldots \tag{A4}$$

Let us suppose $\mathbf{O}(t) = \exp\Omega(t)$, then

$$\frac{d\mathbf{O}}{dt} = \left(\left(\frac{d\Omega}{dt}\frac{\partial}{\partial\Omega}\right)\mathrm{e}^{\Omega}\right)\mathrm{e}^{-\Omega}\mathrm{e}^{\Omega}.$$

By using Formula 2 with $y = \frac{d\Omega}{dt}$ and $x = \Omega$, one obtains

$$\frac{d\mathbf{O}}{dt} = \left\{\frac{d\Omega}{dt}, \frac{1-\mathrm{e}^{-\Omega}}{\Omega}\right\}\mathrm{e}^{\Omega},$$

which compared with Eq. (A3) gives

$$\mathbf{K}(t) = \left\{\frac{d\Omega}{dt}, \frac{1-\mathrm{e}^{-\Omega}}{\Omega}\right\}.$$

We can now apply the Inversion Lemma with the same substitutions in x and y as before, and with $u = \mathbf{K}$ and

$$P(\Omega) = \frac{1-\mathrm{e}^{-\Omega}}{\Omega},$$

to obtain

$$Q(\Omega) = \frac{\Omega}{1-\mathrm{e}^{-\Omega}}$$

and

$$\frac{d\Omega}{dt} = \left\{\mathbf{K}, \frac{\Omega}{1-\mathrm{e}^{-\Omega}}\right\}.$$

This expression can be finally expanded using the following power series

$$\frac{\Omega}{1-\mathrm{e}^{-\Omega}} = \sum_{n=0}^{\infty}\beta_n\Omega^n$$

where the β_n has the given values.

To integrate the resulting equation for Ω we start at t=0, where $\Omega(0) = 0$ to satisfy boundary conditions. Introducing this solution into the equation we obtain a new solution:

$$\Omega_1 = \int_0^t \mathbf{K}d\tau.$$

We can iterate the procedure to obtain $\Omega_2, \Omega_3, \ldots$ as

$$\frac{d\Omega_m}{dt} = \sum_{n=0}^{\infty}\beta_n\left\{\mathbf{K}, \Omega_{m-1}^n\right\}.$$

The solution is obtained as the limit of this series:

$$\Omega(t) = \lim_{m\to\infty}\Omega_m.$$

References

van Ballegooijen, A.: 1985, in M. Hagyard (ed.), *Measurements of Solar Vector Magnetic Fields*, p. 322, NASA CP-2374

Beckers, J.M. and Schröter, E.H.: 1969, Sol. Phys. **10**, 384

Collados, M., Martínez Pillet, V., Ruiz Cobo, B., del Toro Iniesta,J.C., and Vázquez, M.: 1994, A&A **291**, 622

Kawakami, H.: 1983, Publ. Astron. Soc. Jpn. **35**, 459

Kjeldseth Moe, O.: 1968, Sol. Phys. **4**, 267

Landi Degl'Innocenti, E. and Landi Degl'Innocenti, M.: 1985, Sol. Phys. **97**, 239

Landi Degl'Innocenti, E.: 1976, A&AS **25**, 379

Landi Degl'Innocenti, E. and Landi Degl'Innocenti, M.: 1981, Nuovo Cimento B **62(1)**, 1

Magnus, W.: 1954, Comm. Pure. App. Math. **VII**, 649

Rachkowsky, D.: 1962, Izv. Krym. Astrofiz. Obs. **27**, 148

Rachkowsky, D.: 1967, Izv. Krym. Astrofiz. Obs. **37**, 56

Rees, D.: 1987, in W. Kalkofen (ed.), *Numerical Radiative Transfer*, p. 213

Rees, D., Murphy, G., and Durrant, C.: 1989, AJ **339**, 1093

Ruiz Cobo, B. and del Toro Iniesta, J.C.: 1994, A&A **283**, 129

Ruiz Cobo, B. and del Toro Iniesta,J.C.: 1992, AJ **398**, 375

Stenflo, J.: 1971, in R. Howard (ed.), *Solar Magnetic Fields*, IAU Symp. 43, p. 101

Stenflo, J.: 1994, *Solar Magnetic Fields*, Kluwer

del Toro Iniesta, J.C. and Ruiz Cobo, B.: 1996, in N. Mein and S. Sahal-Bréchot (eds.), Science with THEMIS, p. 93, Observatoire de Paris-Meudon

del Toro Iniesta,J.C. and Ruiz Cobo, B.: 1996, Sol. Phys. **164**, 169

Unno, W.: 1956, Publ. Astron. Soc. Jpn. **8**, 108

Westendorp Plaza, C., del Toro Iniesta, J.C., Ruiz Cobo, B., Martínez Pillet, V., Lites, B., and Skumanich, A.: 1997a, in B. Schmieder, del Toro Iniesta, J.C., and M. Vázquez (eds.), 1^{st} *Advances in solar physics euroconference. Advances in the physics of sunspots*, Vol. 118 of *ASP Conference Series*, p. 197

Westendorp Plaza, C., del Toro Iniesta, J.C., Ruiz Cobo, B., Martínez Pillet, V., Lites, B., and Skumanich, A.: 1997b, in B. Schmieder, del Toro Iniesta, J.C., and M. Vázquez (eds.), 1^{st} *Advances in solar physics euroconference. Advances in the physics of sunspots*, Vol. 118 of *ASP Conference Series*, p. 202

Westendorp Plaza, C., del Toro Iniesta, J.C., Ruiz Cobo, B., Martínez Pillet, V., Lites, B., and Skumanich, A.: 1997c, Nature **389**, 47

Wittmann, A.: 1974, Sol. Phys. **35**, 11